PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : G01S 3/02

A1

(11) International Publication Number: WO 00/34799

(43) International Publication Date: 15 June 2000 (15.06.00)

(21) International Application Number: PCT/US99/24493

(22) International Filing Date: 19 October 1999 (19.10.99)

(30) Priority Data:

09/205,557	4 December 1998 (04.12.98)	US
09/231,256	15 January 1999 (15.01.99)	US
09/275,655	24 March 1999 (24.03.99)	US

(71) Applicant *(for all designated States except US)*: U.S. WIRELESS CORPORATION [US/US]; Suite 200, 2303 Camino Ramon, San Ramon, CA 95483 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: HILSENRATH, Oliver [–/US]; 32 Essex Court, Alamo, CA 94507 (US). WAX, Mati [–/US]; Apartment 26, 120 Reflections Drive, San Ramon, CA 94583 (US). JAYARAMAN, Sriram [–/US]; 3300 Treetops Circle, #201, San Bruno, CA 94066 (US). BAR, Abraham [–/US]; 936 El Cajon Way, Palo Alto, CA 94303 (US). MALKIN, Mathan [–/US]; Apartment 411, 105 Shoreline Circle, San Ramon, CA 94583 (US). RAJAPAKSE, Ravi [–/US]; 1448 Pine Street #304, San Francisco, CA 94109 (US). RADIONOV, Vladimir [–/US]; 130 Reflections Drive, #12, San Ramon, CA 94583 (US). MENG, Yan [–/US]; 275 Reflections Drive, #27, San Ramon, CA 94583 (US).

(74) Agent: ALBOSZTA, Marek; 426 Lowell Avenue, Palo Alto, CA 94301 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

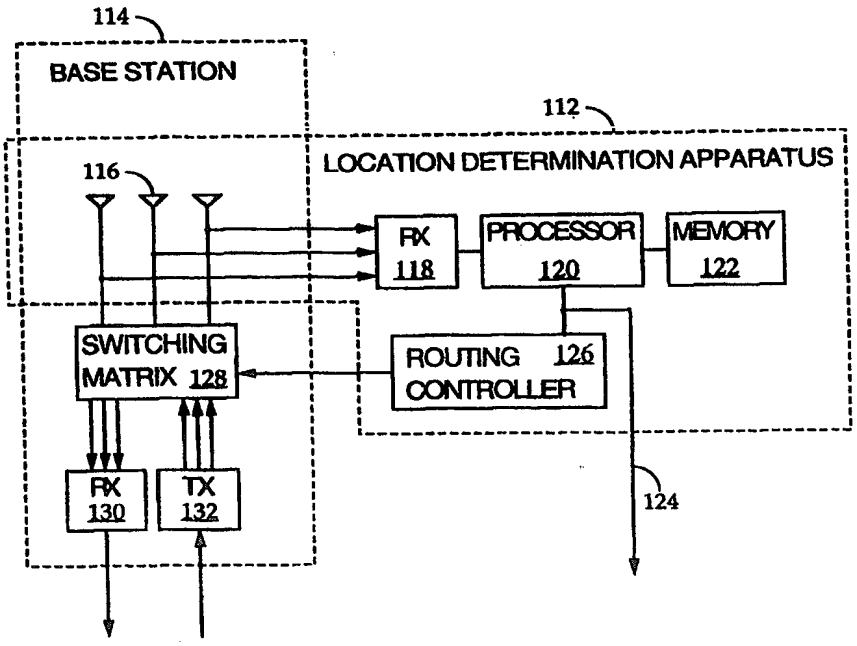
With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: WIRELESS LOCATION DETERMINATION USING SPATIAL SIGNATURE INFORMATION

(57) Abstract

A method for determining locations of mobile transmitters in a wireless communication system calculates a signature from signals received at an antenna array from the mobile, and matches the signature with calibrated signatures stored in a database to estimate the location of the mobile. The signature contains information characterizing the spatial channel between the mobile and the antenna array. An apparatus (112) implementing the method comprises an antenna array (116) and multi-channel receiver (118) for coherently receiving multi-dimensional signal vectors from a mobile transmitter. A processor (120) calculates a signal signature from the received signal vectors, which compared with calibrated signatures stored in a memory (122) to identify calibrated signatures that are similar to the calculated signature, and which corresponds to the likely locations of the mobile transmitter. Thus, the location of the transmitter can be accurately determined from the signals received at a single base station, even in a severe multipath environment.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

BNSDOCID <WO_____0034799A1 I >

PCT PATENT APPLICATION OF U.S. WIRELESS CORPORATION

WIRELESS LOCATION DETERMINATION USING SPATIAL SIGNATURE INFORMATION

FIELD OF THE INVENTION

This invention relates generally to methods for determining the location of a radio transmitter. More particularly, it relates to methods for determining the location of cellular telephones, pagers, and other portable transmitters operating in urban environments characterized by a high degree of multipath.

BACKGROUND OF THE INVENTION

A long standing problem in radio communications is accurately determining the location of a mobile radio transmitter. Precise location information in a cellular telephone network is desirable for various reasons that have been recognized in the prior art. For example, cellular location information can be used to assist emergency dispatching, tracking unauthorized cell phone usage, and tracking or locating fleets of vehicles. Other applications include locating lost or stolen vehicles, assisting lost motorists, and providing various location-based services. These applications have yet to be realized, however, because they require accurate location information and no suitable method of providing such information has yet been developed. Generally, prior art location determination techniques fail for one or more of the following reasons: (1) they fail to operate in urban environments where multipath effects are severe, (2) they do not operate if fewer than three base stations are in range of the signal from a transmitter, (3) they do not provide location information with sufficient accuracy, precision, or confidence, and (4) they require extensive modifications to the mobile transmitter, to the communication system base stations, or to both.

Multipath is perhaps the most fundamental obstacle to prior art methods of solving the location finding problem. Multipath is typically caused by the reflection of signals from objects in the environment, such as buildings, hills, and other structures. A signal transmitted from a cellular phone will often be reflected from these structures, resulting in

multipath signals arriving at a base station in addition to a direct path signal. The angles of arrival of the multipath signals are usually different from the angle of arrival of the direct path signal. In addition, because the path lengths of the multipath signals differ from those of the direct path signal, the multipath signals and the direct path signal have different times of arrival. Moreover, in an urban environment where severe multipath is present, sometimes no direct path exists at all and the base station receives only the multipath signals. Conventional location determination techniques, however, are based on angle-of-arrival measurements or time-of-arrival measurements. As a result, they suffer significantly from the effects of multipath, and elaborate techniques are used in attempts to mitigate the multipath effects. For example, the prior art teaches signal averaging techniques, spread spectrum or frequency hopping techniques, and techniques for multipath identification and removal. In cases of severe multipath, however, there often is no direct path signal at all. In such cases, these approaches fail. Moreover, averaging techniques are based on assumptions about the distribution of multipath that are not generally valid, especially in severe multipath environments. Even in cases where such assumptions do hold, these averaging techniques do not yield accurate position information. The prior art approaches, therefore, operate only when multipath effects are minimal, and fail to provide reliable and accurate location information in urban environments where multipath effects are severe.

SUMMARY OF THE INVENTION

In contrast with all prior art location finding methods which attempt to eliminate or avoid multipath signals and their effects, the present invention advantageously uses multipath signals in order to accurately determine a transmitter's location. Surprisingly, in contrast to other techniques of location finding, the accuracy of the present technique is improved in severe multipath environments. Moreover, the present invention provides a method of location finding which, in contrast to other methods, can accurately determine a location from a single base station. This feature of the invention also distinguishes it from prior techniques which either require multiple base stations or provide only very inaccurate or imprecise location information. In addition, the present location determination technique does not require any modification to the transmitter, or to the existing base stations of the wireless communication system. The invention, therefore, provides a unique solution to the location finding problem which has heretofore remained unsolved.

According to the principles of the present invention, an antenna array of a location-finding apparatus coherently receives multi-dimensional signal vectors from a mobile transmitter. The specific technique for acquiring the signal vectors will depend on the wireless communication standard used by the transmitter. Once the signals are acquired, the apparatus

BNSDOCID: <WO_____0034799A1_I_>

then calculates a transmitter signal signature from the received signal vectors. In general, a signal signature is defined as any location-dependent feature derived from the signals received at the antenna array from a transmitter at a given location. For example, the signature may be derived from various combinations of signal amplitude, phase, delay, direction, and polarization information derived from the signal vectors. Preferably, the signature is substantially invariant with respect to all variables unrelated to transmitter position, such as noise and interference. In a preferred embodiment, the signal signature comprises signal multipath characteristics of a spatial channel between the mobile and the location-finding apparatus. Most preferably, the signature comprises a signal covariance matrix formed from the signal vectors, or a subspace of the matrix spanned by its dominant eigenvectors.

The transmitter signal signature, together with calibrated signatures and corresponding calibrated locations are then used to determine one or more likely locations of the transmitter, for example, by calculating a similarity between the transmitter signature and each of the calibrated signatures. The likely locations correspond to the calibrated signatures that are most similar to the transmitter signature. In this manner, the location of the transmitter can be accurately determined from signals received at a single base station, even in a severe multipath environment. More sophisticated matching techniques can be used to reduce location ambiguity and provide more accurate location estimates. In addition, although the present techniques do not require more than one base station for accurate location determination, in the event that signatures from multiple base stations are available, they can be combined to increase the accuracy and confidence of the location estimates.

The calibrated signal signatures and corresponding locations are determined by a calibration procedure in which known locations are associated with transmitter signal signatures, preferably using a calibration transmitter equipped with a global positioning system (GPS) receiver. The calibrated signature and location information can then be collected, processed, organized, and stored in a database for subsequent use in real-time location determination.

The techniques of the present invention have various useful applications. For example, the database of signal signatures and locations can be used to derive antenna calibration information. This information, in turn, can be used to enhance location determination. In another useful application, the signature of a transmitter can be used to estimate its velocity. This velocity information, in turn, can be used to improve signal acquisition and signature determination, as well as location determination. The accurate location information provided by the present invention has a broad range of useful applications. To name just a few, the location information can be used to provide location-based services and information to mobile users, it can be used to track or locate people, vehicles, or other valuable objects that

BNSDOCID: <WO_____0034799A1_I_>

may be missing, stolen, or in need of attention or assistance, and it can be used to provide to a wireless communications system operator real-time location-enhanced information that can assist in various types of network enhancement, diagnosis, planning, evaluation, analysis, or management.

DESCRIPTION OF THE DRAWINGS

FIGS. 1A, 1B, and 1C illustrate three levels of integration between an existing base station and a location determination apparatus (LDA) of the present invention. FIG. 1A shows an LDA closely integrated with a base station, FIG. 1B shows an LDA partially integrated with a base station, and FIG. 1C shows a stand-alone implementation of an LDA not directly connected to the base station.FIG. 2 is a block diagram of an LDA integrated with a base station, corresponding to the implementation shown in FIG. 1A.

FIG. 3 illustrates a spatial channel that is characterized by multipath propagation effects. According to the present invention, this spatial channel information measured and used to construct a signature that is highly correlated with the mobile position. Comparing the signature with calibrated signatures associated with known locations, the location of the mobile is estimated.

FIG. 4 is a block diagram of an LDA according to an embodiment of the present invention suited for a CDMA wireless system.

FIG. 5 illustrates the temporal structure of a CDMA signal according to the IS-95 standard. In particular, the CDMA signal is composed of a sequence of temporal frames, each having a duration of 20 ms. Each frame, in turn, is divided into 16 power control groups, each having a 1.25 ms duration.

FIG. 6 is a block diagram of an efficient receiver architecture according to an embodiment of the present invention suited for a CDMA wireless system.

FIGS. 7A, 7B, 7C, and 7D illustrate signal sampling schemes employed in a preferred embodiment of the present invention. FIG. 7A illustrates how approximately 100 signal samples are taken during a 1 ms sampling interval. FIG. 7B illustrates how a batch of signal samples from a given mobile device is collected once every 50 ms. FIG. 7C illustrates how batch covariance matrices for a given mobile device are collected over the course of a few seconds. FIG. 7D illustrates how a sequence of signal covariance matrices is generated over the course of time, with one covariance matrix added to the sequence every 5 s.

FIGS. 8A and 8B are schematic diagrams of geographical bins and sub-bins used in a calibration table compression technique of the present invention. The sub-bins shown

BNSDOCID <WO_____0034799A1_I_>

are squares about 10 m across, and the bins are squares about 30 m across. The locations of raw data points are shown in the figures as empty circles.

FIG. 9 is an illustration of the service region of an LDA containing a set of calibrated locations, shown in the figure as circled 'x' marks. The figure also shows an angular sector of the service region centered around an angle θ having an angular width $\Delta\theta$. According to one aspect of the invention, the calibrated signatures in the sector are used to calculate an array calibration vector.

FIG. 10 is a graph of a function $D(S,S_k)$ with respect to calibrated locations k. The value of the function at a given location k is a measure of the similarity between a measured signature S and a calibrated signature S_k. The graph also represents a "spectrum" of the signature S with respect to the entire set of calibrated signatures $S_1,\ldots,S_N$. According to one matching technique, the likely locations of the mobile having signature S are the locations $k(1),k(2),\ldots,k(M)$ such that $D(S,S_{k(j)}) > h$ for $j=1,\ldots M,,$ where h is a predetermined threshold value. This subset of points is illustrated in the figure as those points above the dotted line corresponding to the value of h.

FIG. 11 shows the graph for a measured signal spectrum w together with the graphs for two calibrated spectra $w_{k(1)}$ and $w_{k(2)}$. According to a preferred matching technique, the likely locations of a mobile having spectrum w are found by comparing w with the calibrated spectra $w_1,\ldots,w_N$. For example, as is evident from the figure, the spectrum $w_{k(2)}$ is more similar to the spectrum w than to the spectrum $w_{k(1)}$. Thus, location k(2) is more likely than location k(1).

FIG. 12 is a graphical representation of mobile "tracks" according to a preferred ambiguity reduction technique. The tracks are illustrated in the figure by solid curves. The tracks, which represent past hypothetical trajectories derived from sets of likely locations collected over time, are used to predict a set of next points, shown as empty circles. Ambiguities in a current set of likely locations can be reduced by reducing the probabilities for locations that are not close to the predicted next points.

FIG. 13 illustrates a system used to implement a technique for resolving ambiguity by combining sets of likely locations obtained from multiple LDAs. A central hub collects the location information over communication lines. Because the location information is independently derived at separate LDAs, it can be combined to reduce ambiguity and increase confidence.

FIG. 14 is a schematic diagram of a hierarchical tree structure used in a technique to improve the computational efficiency of the real-time signature matching. The tree structure is composed of a number of nodes, shown as circles in the figure. Each calibrated location is assigned to at least one bottom node. The nodes higher in the tree contain the union of all nodes connected downward from them. In the search of similar

BNSDOCID: <WO_____0034799A1_I_>

calibrated signatures, this tree can be used to eliminate a significant amount of computation.

DETAILED DESCRIPTION

Although the following detailed description contains many specifics for the purposes of illustration, anyone of ordinary skill in the art will appreciate that many variations and alterations to the following details are within the scope of the invention. Accordingly, the following preferred embodiment of the invention is set forth without any loss of generality to, and without imposing limitations upon, the claimed invention.

System Architecture

NETWORK OVERVIEW

A preferred embodiment of the present invention is designed to operate in the context of one or more wireless communication systems providing services to cellular telephones, pagers, wireless modems, and other portable communication devices. Each communication system includes a number of base stations, each designed to communicate with the portable devices in its geographical area, or cell. Downlink radio frequency (RF) signals are transmitted from the base station to the portable devices, and uplink RF signals are transmitted from the portable devices to the base station. A conventional base station typically has a set of six directional antennas for transmitting the downlink signals and receiving the uplink signals. These six antennas are normally grouped into three pairs, where each pair is designed to serve a 120-degree sector of the cell. When communicating with a portable device in a particular sector, only one of the six antennas is normally used. In particular, the antenna used for a particular mobile device is dynamically selected to provide the best signal. If the quality of the communication link between a base station and a mobile device degrades, the communication system typically attempts to hand-off the link to another base station that can provide a higher quality link. In this manner, a portable device moving throughout the communication system's network of cells can receive continuous service. The location of a particular device in these conventional systems is known only approximately. The network knows which base station is serving the mobile device, and can sometimes determine which antenna is being used for communication with the device. Using this information, however, the location of a user can be narrowed down to an area of several square kilometers, at best.

BNSDOCID: <WO_____0034799A1 I_>

LOCATION FINDING APPARATUS

A location determination apparatus (LDA) according to the present invention is used to provide more accurate and precise location information of mobile devices. As illustrated in FIGS. 1A, 1B, and 1C, the LDA may be integrated partly or completely with an existing base station of a communication system, or may be a separate stand-alone apparatus operating idependent of other communication systems. For example, FIG. 1A shows and LDA 100 integrated with a base station 102. In this integrated implementation, the LFA makes use of the existing antennas of the base station, and has no antennas of its own. Because these antennas are directly connected to the LDA, they are considered part of the LDA. In a highly integrated implementation, the LDA is not only connected to the base station antennas, but is also connected to the internal electronic equipment of the base station. Such a connection can be used, for example, to communicate user channel assignment information from the base station to the LDA, and/or to communicate location information from the LDA to the base station. In a partial integration, as shown in FIG. 1B, an LDA 104 has its own antennas and is independent and separate from the base station, with the exception of a communication link for exchanging information. In a stand-alone implementation, shown in FIG. 1C, an LFA 108 and a base station 110 are not directly connected at all, although they may be indirectly connected via a larger network.

An integrated implementation of an LDA 112 with a base station 114 is illustrated in FIG. 2. The LDA 112 comprises a set of p antennas 116, a multichannel receiver 118 which coherently receives signals from the p antennas, a multi-purpose signal processor 120 which performs signal signature identification, database searching, and other required processing as will be described in detail below, and a memory storage device 122 accessible by the processor and used to store signal signature calibration data, location data, geographical map data, and/or other data as required. Location information determined by the processor is preferably sent over external communication lines 124 for use by other systems. The location information may also be provided to a routing controller 126 connected to a switching matrix 128 of the base station 114. The switching matrix 128 performs antenna beamforming or switching in dependence upon location-dependent signals from routing controller 126. A receiver 130 and transmitter 132 of the base station are connected to the antennas 116 by means of the switching matrix. Thus, the transmission and reception of the base station may be controlled by location-based routing information derived from LDA 112. The implementation shown in FIG. 2 is one example of a highly integrated implementation, corresponding to FIG. 1A. A stand-alone implementation of LDA 112 would not have routing controller 126, and would not share antennas 116 with base station 114. The essential principles of operation would otherwise be the same.

BNSDOCID: <WO_____0034799A1_I_>

In contrast with the base station which communicates with a particular mobile using just one selected antenna at any given time, the LDA coherently receives signals from all the antennas simultaneously. The coherent multichannel receivers 118 preserve relative phase information between the signals received at the various antennas 116. As a result, subtle information about the spatial channel can be determined, as will be described in detail below. In particular, this spatial channel information includes multipath signal information, as shown in FIG. 3. Typically, in addition to a direct path signal 134 from a mobile device 136 to an LDA 138, there are one or more additional multipath signals 140 reflected from various environmental objects. In addition, there are typically various smaller local scatterers (not shown) positioned within a region 142 around the device. The coherently received signals at the LDA 138 thus include subtle spatial channel information corresponding to these reflected and scattered signals.

ARCHITECTURE FOR CDMA SYSTEMS

The wireless communication system can employ various conventional air standards and protocols, such as AMPS and CDMA. In the case of the more complicated CDMA standard, such as IS-95, some additional description will be helpful. In a CDMA communication system, all user devices share a common frequency band, and the different mobile devices are distinguished by their use of distinct PN spreading sequences. At the base station, the signals from a device are coupled to one of the antennas, downconverted to IF, and then digitized. The digitized signals are then correlated with PN sequences in order to separate the signals from different mobile devices. Low cross-correlation between the unique PN sequences ensures that there is little interference between mobiles within the base station service area. Well-known timing synchronization techniques are used to identify and lock onto the PN sequence for each mobile. After the user signal has been separated, it is typically sent into an inverse Hadamard transformer in order to extract Walsh symbols from the signal. Various decoding and interleaving functions may also be performed in order to reconstruct the information signal originating at the mobile device, as is well known in the art.

In an LDA 144 according to a CDMA embodiment of the present invention, the p antenna signals are coupled to coherent multichannel IF down-converters and A/D converters 146 to produce a set of p digitized complex signals, as shown in FIG. 4. Down-conversion and digitization are performed coherently so as to preserve the phase relationships between the separate signals from the p antennas. A common local oscillator and a master clock are used to maintain coherence between the parallel functions performed on all the signals from the p antennas. In the figure, the LDA 144 is integrated with a base station 148 which has its own

BNSDOCID. <WO_____0034799A1_I_>

transmitter 150 and receiver 152. Stand-alone and partially integrated implementations, however, are also possible.

The p complex signals are separately despread in synchronized multichannel despreaders 154 of a CDMA receiver 156 of the LDA. In one embodiment, user PN sequence information is provided from the base station 148 so that the LDA can despread the signals. Alternatively, in stand-alone implementations, the LDA can obtain this information from the user device. For example, the user device can transmit its code channel information to an internet computer via a wireless modem link. The LDA can then query the computer and obtain the user information, thus by-passing the need to obtain it from the base station. In any case, once the PN sequence information for the user is obtained, the LDA despreads the p signals. In contrast with the base station receiver 152 which simultaneously performs despreading for all mobiles being served, the receiver 156 of the LDA sequentially time samples the signals from the set of current mobiles, thereby dramatically reducing hardware requirements and cost. To perform this sequential despreading, accurate code timing information is required. This information can be obtained from base station receiver 152. Alternatively, the code timing information can be acquired and tracked by the receiver of the LDA using well known techniques. Intermittent time samples are sufficient for the LDA because its purpose is not to acquire all the original signal information, but to acquire a spatial channel signature.

Although the LDA receiver architecture in a CDMA system is preferably adapted in the above manner, the high-level LDA architecture and principle of operation is the same.

Signal Measurement

Signal Vector Sample Acquisition

The LDA hardware coherently receives complex signal samples. These samples are later used to calculate spatial channel signatures, which are the basis for location determination. More specifically, the p antennas are used to receive complex signals $x_1(t)$, $x_2(t)$,..., $x_p(t)$, respectively, which are represented by a p-dimensional signal vector $\mathbf{x}(t) = [x_1(t), x_2(t),..., x_p(t)]^T$. At a moment t in time, each component $x_j(t)$ of this vector is a complex number having both amplitude and phase information from one of the antennas. In general, the dimension of the array vector $\mathbf{x}(t)$ is equal to p, the number of antennas in the array. The value of p is preferably in the range from 2 to 12, and is more preferably in the range from 4 to 8. Most preferably, the value of p is conveniently chosen to be 6 so that the existing antennas of a base station may be used.

BNSDOCID: <WO_____0034799A1_I_>

SIGNAL ACQUISITION IN CDMA SYSTEMS

Signal acquisition in a CDMA system is more complicated, and therefore will be discussed in more detail. As is well known in the art, a conventional CDMA receiver can isolate multipath parts of a signal if their differential delay is more than one chip length. The signal parts are normally extracted by "fingers" of the receiver and then combined after temporal alignment to get improved signal strength. Each finger comprises a correlation receiver matched to the exact delay associated with a particular multipath part. Note, however, that this standard technique only distinguishes temporally distinct multipath parts, and does not distinguish spatially distinct multipath parts that have similar temporal delays.

The CDMA receiver 156 in the LDA 144 of FIG. 4 adapts the above technique for separating multipath parts in order to provide refined discrimination in location finding. For simplicity in the following discussion, we consider a signal from a single mobile device. At the p antennas of the LDA, this signal gives rise to a set of p coherently received signals $x_1(t)$, $x_2(t)$,..., $x_p(t)$. In the synchronized multichannel despreaders 154, each component $x_j(t)$ of this set is then decomposed using q correlation receivers into q temporally distinct signals $x_{j,1}(t),...,x_{j,q}(t)$, where j is the antenna index of the signal and q is the number of fingers used. Thus, a total of p×q correlation receivers are used. Preferably, q is in the range from 1 to 8. More preferably q is in the range from 2 to 6. Most preferably, q=4. The multipath parts contained within a given finger all have delays within one chip from each other, while the delays between fingers are more than one chip. Each set of p samples from a given finger i is represented by a p-dimensional vector $\mathbf{x}^{(i)}(t)=[x_{1,i}(t),...,x_{p,i}(t)]^T$. All q vectors $\mathbf{x}^{(1)},...,\mathbf{x}^{(q)}$ are then sent to a signal processor 158 for subspace determination, as will be described further below.

The CDMA receiver 156 contains a searcher 160 that searches the original signals $x_1(t)$, $x_2(t)$,..., $x_p(t)$ for the best code delays to be used by the q despreaders. The searcher produces a code correlation output that represents the degree to which the delayed code matches the incoming signals. A sequence of L delay values $\tau_1,...,\tau_L$ are selected that are spaced by 1/2 chip length. Typically, L is approximately 200, so that the temporal span of the sequence is 100 chips. At each of the delay values, a corresponding code correlation value for the signal is measured from the code correlation output. The code correlation values from the j^{th} antenna is represented as a sequence $C^{(j)}(\tau_1),...,C^{(j)}(\tau_L)$, where $C^{(j)}(\tau_i)$ is the value of the correlation between the signal $x_j(t)$ and a code with temporal delay τ_i. The

BNSDOCID: <WO_____0034799A1_I_>

correlation output sequences for all p signals are sent to the signal processor 158 for construction of the correlation signature.

The searcher 160 examines the code correlation output to determine a set of q time delays $\tau^{(1)},...,\tau^{(q)}$ which correspond to the strongest correlations. There are various well known methods of determining such time delays. For example, these time delays may be determined by selecting the q values of τ that correspond to the q largest local maxima in the set $\{C^{(j)}(\tau_i) : j=1,...p; i=1,...,L\}$. Once these delays are determined, they are then sent to the despreaders 154 to control the despreading of each of the q fingers corresponding to these delays. The delays are also sent to the processor 158 for signature determination.

In order to achieve optimal use of hardware, the signals from the receivers can be sampled at regular time intervals. This optimization, however, requires some modification of conventional CDMA receiver architecture, which will now be described in detail with reference to FIGS. 5 and 6. In the following description, it is assumed that p=6 and that the IS-95 standard is used. The principles of this optimization, however, do not depend on these particular assumptions. Moreover, it should be emphasized that the present invention does not require this optimization in order to be usefully implemented.

It should be noted that the CDMA signal from each device has a standard temporal structure. In particular, the CDMA signal is composed of a sequence of temporal frames, each having a duration of 20 ms and containing 192 bits of information, as shown in FIG. 5. Each frame, in turn, is divided into 16 power control groups, each having a 1.25 ms duration. At high data rates, all 16 power control groups contain information. At the lowest data rate, only two power groups are active, i.e. contain information. At low rates, the placement of the active groups within the frame is pseudorandom.

As shown in FIG. 6, the signals from each user device in the coverage area of the LDA are coupled to the six antennas and received by a corresponding bank of six receivers 164. These receivers are coherently synchronized so that they preserve the relative phase information between the signals coupled to the six antennas. Each receiver downconverts the signal from one antenna to baseband, and then digitizes the signal to produce I and Q signal components. At this point, the signals from the various user devices have not yet been distinguished. The six digitized signals produced by the receiver bank 164 are sent to a corresponding bank of six signal buffers 166. Each signal buffer stores data samples covering several recently received digitized signal frames. Typically, each buffer is large enough to store three or more frames of data. At periodic intervals of 1.25 ms, a buffer

BNSDOCID: <WO_____0034799A1_I_>

control circuit 168 calculates a buffer address that is used to select a power control group from the signal frames currently in each buffer. The specific buffer address is calculated so that during each power group time interval an active power group is selected from one of up to 16 distinct user phones. The output from each of the six buffers, therefore, is a data stream of selected power groups delayed by approximately three frames, where each of the 16 power group intervals in a frame can contain an active power group from one of up to 16 distinct user signals. Buffer control 168 sends a common set of signals to the buffers in signal buffer bank 166. These signals include buffer record and playback addresses, read/write signals, and buffer configuration data.

The six data streams from the buffer bank 166 are sent for despreading to a bank of six corresponding despreaders 170. In each despreader, the incoming data stream of selected power groups is correlated with appropriate PN sequences in order to despread the data. In a CDMA communication system, all users share a common frequency band, and the different mobile phones are distinguished by their use of distinct PN spreading sequences. Low cross-correlation between the unique PN sequences ensures that there is little interference between different users within the base station service area. In addition, low cross-correlation between distinct delayed versions of a given PN sequence ensures that there is little interference between various multipath parts of a signal originating from a single user. Well-known timing synchronization techniques are used to identify and lock onto the PN sequences of the multipath parts for each user.

In contrast with conventional use, despreaders in the present apparatus have their PN sequences changed every 1.25 ms by a despreader control circuit 172. The PN sequence is changed in synchronization with the transition between power groups so that, during each of the 16 power groups in a frame, the despreader uses the PN sequence corresponding to one of 16 users. The output from each despreader, therefore, is a despreaded I/Q data stream where each of the 16 power group intervals in a frame contains the despreaded signal from one of up to 16 user signals.

It is significant to note that the resulting six signal data streams do not contain all the information transmitted from the mobile phones. Rather, each frame coming out of the despreaders contains up to 16 active power group samples derived from up to 16 different mobile phones. Note that since the despreading essentially separates the multipath parts with differential delays greater than the chip period, it is possible to obtain the spatial signatures of all the significant multipath parts (referred to as fingers) of every mobile. In the present system, up to four fingers per mobile can be despread. The despreader output is used for determining spatial signal information, as will be described in more detail below.

BNSDOCID: <WO 0034799A1 I_>

We now turn to a description of the operation of buffer control circuit 168 and despreader control circuit 172. A common set of signals are sent from despreader control 172 to despreader bank 170. Specifically, these signals include two PN sequences: a combined I despreading sequence and a combined Q despreading sequence. Each of these sequences is a chip-by-chip product of all the spreading signals involved in spreading the signal in the transmitter. In the case of IS-95, these spreading signals include the long code, the Walsh codes, and the short I or Q sequences.

It is important to note that the Walsh codes are normally used as the transmitted symbols and thus contain the transmitted signal information. The present technique, however, treats the Walsh code as just another spreading signal. Including the Walsh code as part of the spreading signal provides improvements in processing at the LDA receiver in determining the phase of the received signals. In the preferred embodiment, the Walsh code information is obtained by sending the decoded data bits from the CSM board through a convolutional encoder, an interleaver, and a Walsh encoder, just as was done in the mobile prior to transmission. It will be appreciated that other codes or training sequences used by various other standards can be used just as the Walsh codes are used in the IS-95 standard described in this embodiment. It will also be appreciated that the Walsh codes could, in principle, be provided by the CSM board since the decoders determine the Walsh codes in order to produce the data stream. Although presently existing CSM integrated circuits do not provide Walsh code output, such integrated circuits may be produced in the future and could be used in the context of the present invention.

A despread IRQ (DES. IRQ) signal is issued on any power control group timing interval transition and serves to trigger the microprocessor 174 to load the next set of despreading parameters into the despread control and buffer control. The GPS IRQ is issued on each UTC second boundary and triggers the microprocessor to perform various housekeeping tasks. The system bus 176 carries information between system components. For example, data sent over the bus includes system reset commands, buffer address information for data playback, PN sequence states for despreading, PN mask information for data despreading, Walsh codes for data despreading, and CSM data and control information.

The six digitized signals produced by the receiver bank 164 are sent to a conventional CDMA modem board 178 which contains 16 conventional CSM integrated circuits which contain searchers, despreaders, fast Hadamard transform (FHT) processors, deinterleavers, decoders, and other well-known circuitry required to produce Walsh symbols and other signal information from the signals received from up to 16 mobile devices. The conventional CDMA modem board 178 does not produce spatial signal information. It does, however, provide data transmitted by each one of the user devices, the rates of transmission for each

BNSDOCID: <WO_____0034799A1_I_>

user, PN sequence information, finger offset information for up to eight paths per user, frame offset information, and power group placement information. This information is made available over the data bus 176 to the microprocessor 174, which computes buffer addresses that are sent over the bus to the buffer control circuit 168, and PN sequences that are sent over the bus to the despreader control circuit 172. The purpose of the three-frame delay of the buffers is to allow time for the microprocessor to compute the buffer addresses, PN sequences, and other control parameters. A GPS timing circuit 180 is used to provide precise timing and synchronization information to CDMA modem board 178, buffer control circuit 168, and despreader control circuit 172. The signals sent from GPS timing circuit to receiver board 178, buffer control 168, and despread control 172 include a CHIPx16 clock (1.2288 x 16 MHz) and a pulse per 2 second signal (even second pulse).

The six output data streams from the despreaders 170 are low-pass filtered and then made available on a DSP data bus 182. These samples are then collected by a digital signal processor (DSP) 184. The DSP first performs a frequency estimation that is aimed at nulling the frequency error of the sampled signals. This frequency error is the result of the difference in frequencies between the user transmitter and the base station receiver. Since the frequency error can be up to several hundred Hertz, when the signal is despreaded and downconverted to baseband, it will have a residual non-zero frequency. Therefore, since signal samples in the present system are averaged over relatively long periods of time, the residual frequency must be corrected to prevent the samples from averaging out to zero. The DSP therefore first determines the residual frequency, and then corrects it before averaging the samples. This correction data is used by receiver control 186 to set receiver frequency and attenuation of receivers 164.

The specific functions performed by the microprocessor 174 include:

1. Initializing and controlling CSMs, GPS timing, buffer control, and despreader control.
2. Reading rates, offsets, active power groups, and transmitted data of subscribers from CSMs.
3. Computing transmitted Walsh codes from the received data.
4. Loading Walsh codes into despreader control before each power control group interval transition.
5. Loading playback address into buffer control before each power control group interval transition.
6. Receiving subscriber data from the DSP for despreading.
7. Sending, for each batch of data sent to the DSP from the despreader, a descriptor to the DSP identifying the subscriber associated with the data.

BNSDOCID <WO_____0034799A1_I_>

The specific functions performed by the DSP 184 include:

1. Computing a frequency error for each subscriber.
2. Correcting the frequency error for each received data batch.
3. Computing covariance matrices for each data batch.
4. Transferring covariance matrices to the location finding processor 188.
5. Transferring instructions to the microprocessor (e.g., subscriber lists and parameters)

The specific functions performed by the receiver control 186 include:

1. Setting the receiver frequency.
2. Setting the receiver attenuation.

In order to compute the spatial signature of a CDMA signal, one must first despread it. To determine the spatial signature, however, it is not necessary to capture or even compute the actual information data stream. For the purposes of the spatial signature measurement technique of the present invention, the important information is not the information data stream itself, but rather the spatial channel information derived in part from the relative phase information between the antennas of the array. Consequently, the present technique is distinct in motive, method, and result from known CDMA reception techniques. It provides an efficient method and system for simultaneously determining spatial signal information for multiple users in a CDMA communications system.

Data Collection and Averaging

SIGNAL BATCHES AND STANDARD AVERAGING

In a preferred embodiment, a p-dimensional received signal $\mathbf{x}(t)$ for a specific mobile device is regularly sampled during a sampling interval to produce a batch of signal vector samples, $\mathbf{x}(t_1)$, $\mathbf{x}(t_2)$,..., $\mathbf{x}(t_N)$. In one embodiment illustrated in FIG. 7A, approximately 100 samples are taken during a 1 ms sampling interval, i.e., the signal is sampled every 10 μs, so that the difference $\Delta t_i = |t_i - t_{i+1}|$ is 10 μs. The outer products of these 100 vectors in the 1 ms batch are then combined to form a $p \times p$ covariance matrix for the batch.

BNSDOCID: <WO_____0034799A1_I_>

For each mobile device, a batch of signal samples is collected once every 50 ms, as shown in FIG. 7B, and a corresponding batch covariance matrix is calculated. The batch covariance matrices for a given mobile device are then collected over the course of a few seconds (e.g., 5 s), as shown in FIG. 7C. These batch covariance matrices are then averaged to form a p×p signal covariance matrix, **R**, which is normalized such that Tr **R** = 1. This normalized covariance matrix is used as the basis for subsequent calculations, as will be described below.

Because each batch spans only 1 ms, there is sufficient time using this technique to sample signals from up to 50 different user devices in a 50 ms interval. In other words, the signals from a first user device are sampled during a first 1 ms interval, the signals from a second user device are sampled during a second 1 ms interval, and so on. After 50 ms, the signals from the first user device are again sampled.

Those skilled in the art will appreciate that similar sampling method also may be implemented using various other values for the time intervals, number of signal vectors sampled per batch, etc. In addition, one can also use a non-uniform sampling scheme wherein each mobile is sampled at a period inversely proportional to the speed of its physical movement, or according to other criteria. Although the present description is focused upon the processing of the signals from a single user device, in the preferred embodiment, the techniques described herein are applied to signals from a number of users.

It will be recognized by those skilled in the art that although this sampling procedure is appropriate for the AMPS cellular standard it can be adapted to various other modulation schemes as well. In the case of a CDMA cellular network, for example, spread spectrum signals from the antenna array are digitized and passed through synchronized multichannel despreaders to produce the p-dimensional array vectors $\mathbf{x}(t)$ corresponding to each mobile. These vectors can then be sampled and averaged to form a covariance matrix in the manner outlined above. More specific details of such a CDMA implementation will now be described.

In the preferred embodiment, a batch of signals for a given mobile is collected every 60 ms in a time interval of length 1.25 ms. Thus, up to 48 different mobiles can be sampled in the 60 ms interval. Obviously, more mobiles can be handled by either decreasing the 1.25 ms batch interval length, or increasing the 60 ms repetition interval. In addition, the hardware can be duplicated so that several mobiles are processed in parallel rather than sequentially. These various implementations involve consideration of various trade-offs, but do not alter the essential technique of the present invention.

BNSDOCID <WO_____0034799A1_I >

The outer products of the signals in each batch are combined to form a batch covariance matrix. If signals from multiple fingers are available, there will be a batch covariance matrix for each finger. The batch covariance matrices for each finger are then averaged over several seconds to produce a signal covariance matrix **R**. If there are multiple fingers, there will be multiple signal covariance matrices.

The sampled time delays and correlation data are also averaged in an analogous manner over the several second period. These averaging techniques help to reduce the effects of noise in the measured signature.

VELOCITY DEPENDENT AVERAGING

In addition to uniform averaging, various types of nonuniform averaging can be used to increase the efficiency and performance of the signal acquisition. In particular, estimates of the mobile device velocity can be used to control the averaging, as will now be described in relation to FIG. 7D. For a given mobile device, the batch covariance matrices for approximately 100 batches are averaged over a few seconds as described earlier to form a p×p signal covariance matrix **R**. According to the present velocity-dependent averaging technique, this initial covariance matrix is denoted $\mathbf{R}_{t_0}$, where t_0 is the time **R** is measured. As the system continues to operate in real time, another signal covariance matrix $\mathbf{R}_{t_1}$ is determined in the same way from the batch covariance matrices measured over the course of a few seconds that elapse between time t_0 and a later time t_1. In this manner, a sequence of signal covariance matrices $\mathbf{R}_{t_0}$, $\mathbf{R}_{t_1}$, $\mathbf{R}_{t_2}$,..., $\mathbf{R}_{t_n}$ is generated, where $\mathbf{R}_{t_n}$ denotes the most recent covariance matrix. Preferably, the difference $\Delta t_i = |t_i - t_{i+1}|$ is in the range from 0.1 to 20 seconds, but is more preferably between 1 and 10 seconds, and most preferably 5 seconds (assuming uniform sampling).

In accordance with the teaching of the present invention, an average signal covariance $\bar{\mathbf{R}}_{t_n}$ is produced in real time by forming a linear combination of the covariance matrices $\mathbf{R}_{t_0}$, $\mathbf{R}_{t_1}$, $\mathbf{R}_{t_2}$,..., $\mathbf{R}_{t_n}$. In the preferred embodiment, the specific manner in which the covariance matrices are combined to produce the average covariance matrix depends on estimates of the transmitter velocity, represented by $v_{t_0},\ldots,v_{t_n}$. Preferably, the covariance matrices combined to form the average covariance matrix are those collected over a time interval in which the transmitter has moved a distance D, where D is a predetermined distance corresponding to the average spatial span between the calibrated signature locations. Generally, if the velocity estimate is large, then the matrices are combined such that the most

BNSDOCID: <WO_____0034799A1_I_>

recent covariance matrices contribute significantly to the average covariance, while the less recent covariance matrices contribute relatively little to the average covariance matrix. Conversely, if the velocity estimate is small, then the contribution of the less recent covariance matrices is not so small. The velocity estimates can be determined using various techniques known in the art. Alternatively, velocity can be estimated based on recent location estimates derived from techniques of the present invention, or directly from recent covariance matrices, as will be described below. For the purposes of this description, we assume velocity estimates are positive velocity magnitudes.

In one implementation of the preferred embodiment, an average covariance matrix $\bar{\mathbf{R}}_{t_n}$ is calculated by evaluating

$$\bar{\mathbf{R}}_{t_n} = (1/D_m) \sum_{i=m}^{n} v_{t_i} \Delta t_i \mathbf{R}_{t_i} = (v_{t_m} \Delta t_m \mathbf{R}_{t_m} + \ldots + v_{t_n} \Delta t_n \mathbf{R}_{t_n})/D_m,$$

where $D_m = v_{t_m} \Delta t_m + \ldots + v_{t_n} \Delta t_n$ and m is selected such that $D_m \approx D$. In other words, m is selected so that the past covariance matrices $\mathbf{R}_{t_m}, \ldots, \mathbf{R}_{t_n}$ are those collected over a time interval in which the transmitter has moved a distance D. It will be noted that, assuming the time intervals $\Delta t_m, \ldots, \Delta t_n$ are constant, the coefficients that determine the relative weights of the past covariance matrices are proportional to the estimated transmitter velocities. A covariance matrix corresponding to a small velocity will therefore contribute less to the average covariance matrix than a covariance matrix corresponding to a large velocity. In accordance with this general principle, those skilled in the art will be able to construct many other formulas for calculating the average covariance matrix from past covariance matrices.

In another implementation of the preferred embodiment, the average covariance matrix is calculated from the most recent past covariance matrix and velocity estimate and the last average covariance matrix. In other words, $\bar{\mathbf{R}}_{t_n}$ is calculated recursively by evaluating

$$\bar{\mathbf{R}}_{t_n} = f(v_{t_n}, \mathbf{R}_{t_n}, \bar{\mathbf{R}}_{t_{n-1}}),$$

where f is a function of the current estimated velocity of the transmitter v_{t_n}, the current covariance matrix $\mathbf{R}_{t_n}$, and the most recent prior average covariance matrix $\bar{\mathbf{R}}_{t_{n-1}}$. This is a recursive definition since the covariance matrix $\bar{\mathbf{R}}_{t_{n-1}}$ is similarly calculated from

BNSDOCID <WO___ ___0034799A1_I >

$$\bar{\mathbf{R}}_{t_{n-1}} = f(v_{t_{n-1}}, \mathbf{R}_{t_{n-1}}, \bar{\mathbf{R}}_{t_{n-2}}).$$

Thus, one may write

$$\bar{\mathbf{R}}_{t_n} = f(v_{t_n}, \mathbf{R}_{t_n}, f(v_{t_{n-1}}, \mathbf{R}_{t_{n-1}}, \ldots f(v_{t_1}, \mathbf{R}_{t_1}, f(v_{t_0}, \mathbf{R}_{t_0}, \mathbf{R}_{t_0}))\ldots)),$$

where $v_{t_0},\ldots,v_{t_n}$ are the estimated velocities at times $t_0,\ldots,t_n$. Although many possible functions f may be selected for use, in one implementation the function f is defined by forming the following linear combination of $\bar{\mathbf{R}}_{t_{n-1}}$ and $\mathbf{R}_{t_n}$:

$$\bar{\mathbf{R}}_{t_n} = f(v_{t_n}, \mathbf{R}_{t_n}, \bar{\mathbf{R}}_{t_{n-1}}) = \lambda_{t_n} \bar{\mathbf{R}}_{t_{n-1}} + (1-\lambda_{t_n}) \mathbf{R}_{t_n},$$

where λ_{t_n} is a velocity-dependent coefficient that takes values between 0 and 1. Values of λ_{t_n} near 1 result in large retention of past covariance information, while values of λ_{t_n} near 0 result in large forgetting of past covariance information.

In one implementation of the present invention, the values for $\lambda_{t_0},\ldots,\lambda_{t_n}$ are calculated from time-dependent velocity estimates $v_{t_0},\ldots,v_{t_n}$. The value of λ_{t_n} is calculated such that a very large velocity estimate v_{t_n} results in a value of λ_{t_n} near 0, and hence large forgetting of past covariance information, while small velocity estimates result in values of λ_{t_n} near or equal to 1, and hence large or total retention of past covariance information. Although many formulas have these properties, one example of a formula for calculating λ_{t_n} is the following:

$$\lambda_{t_n} = e^{-v_{t_n}\Delta t_n / D},$$

where Δt_n is on the order of a few seconds, D is on the order of a few tens of meters, and v_{t_n} is the current estimate of the transmitter velocity. For example, in the case where $\Delta t_n = 1$ s and D=25 m, a velocity of $v_{t_n} = 0$ will result in $\lambda_{t_n} = 1$. A velocity of

BNSDOCID: <WO_____0034799A1_I_>

v_{t_n} = 25 m/s = 90 km/hr), on the other hand, will result in $\lambda_{t_n} = e^{-1} = 0.368$. It will be appreciated that many other equations for λ_{t_n} may be used to produce substantially equivalent effects. For example,

$$\lambda_{t_n} = 1/(1 + v_{t_n})$$

or

$$\lambda_{t_n} = 1/(1 + (v_{t_n})^2).$$

In an alternate embodiment, the value for λ_{t_n} is selected from a finite set of predetermined values $\lambda_1,\ldots,\lambda_m$. The particular value selected depends on the value of v_{t_n}. For example,

$$\lambda_{t_n} = \begin{cases} 1 \text{ if } v_{t_n} < v_{low} \\ 0 \text{ if } v_{t_n} > v_{high} \\ .5 \text{ otherwise,} \end{cases}$$

where v_{low} and v_{high} are predetermined low-velocity and high-velocity thresholds, respectively.

This technique for calculating an average covariance matrix from past covariance matrices has important advantages. Because the current covariance matrix of a very slowly moving or stationary transmitter has less information than that of a moving transmitter, it correlates weakly with location and introduces ambiguity into the location finding process. The present technique, however, retains past covariance information when the velocity is small or zero, and overcomes this problem. When the transmitter is moving quickly, on the other hand, the current covariance matrix correlates strongly with the current location, while the past covariance matrices correlate strongly with past locations. Because the present technique forgets past covariance information in this case, it provides an automatic and adaptive way to generate an average covariance matrix that is most highly correlated with the current location.

BNSDOCID: <WO 0034799A1_I_>

Signature Computation

The signal signature for a mobile device captures the essential spatial channel information derived from the signals received at the antenna array. Because the signal signature depends on the spatial channel information, it is highly correlated with the spatial location of the mobile. The signal signature preferably includes location-dependent aspects of the spatial channel, and preferably excludes as much as possible location-independent aspects such as noise. More preferably, the signal signature includes signal multipath information resulting from multipath propagation effects in the environment, preferably derived from phase and amplitude information measured at an antenna array. Most preferably, the signal signature includes one or more of the following elements: a spatial subspace of the signal covariance matrix, relative time delays of signal multipath parts, and code correlation information. The signature can also be represented in other ways, such as higher order spectra of the signals. Signal subspaces may be also extracted via other techniques involving cumulants and higher order statistics in conjuction with second order statistics represented by the signal covariance matrix. Cumulants can then be used to recover more phase information than is possible from the covariance matrix alone and in addition possess more robust properties in the presence of additive Gaussian observation noise.

Covariance Matrix Subspace Calculation

In a preferred embodiment, the signal signature includes as its primary component a signal subspace $\mathbf{U}$ spanned by the dominant eigenvectors of the signal covariance matrix $\mathbf{R}$. The dominant eigenvectors are defined as the eigenvectors of $\mathbf{R}$ corresponding to the largest eigenvalues that collectively capture at least certain percentage of the total signal energy. Preferably, this percentage is at least 70%, but is more preferably 90%. It should be emphasized that the signal subspace may be determined by various different criteria as well.

In the case of a CDMA system, q signal subspaces $\mathbf{U}^{(1)},\ldots, \mathbf{U}^{(q)}$ are preferably calculated for the q signal covariance matrices $\mathbf{R}^{(1)},\ldots, \mathbf{R}^{(q)}$ of the q fingers. It should be noted that each multipath signal in a finger i normally contributes an extra dimension to the subspace $\mathbf{U}^{(i)}$. Consequently, the subspace will have an increased dependence upon location in the presence of severe multipath.

Correlation and Time Delay Information

In addition to the signal subspace, differential time delays for the multipath parts can also be used as part of the signal signature. These delays can be used to further enhance the signal

BNSDOCID: <WO_____0034799A1_I_>

signature in order to provide additional location dependence. As a result, the delays provide the system with increased location accuracy, and also can be used to resolve ambiguities in location, such as may arise in the case of a stationary phone using a narrowband modulation scheme. (Note that, in contrast with the time difference data in the prior art which consists of time differences of the direct signals arriving at several different base stations, the differential time delay data used herein involves the time differences between the multipath signal parts arriving at a single base station.)

In one embodiment, differential time delays between multipath signals may be calculated as follows. First one performs a blind estimation of the array response matrix **A** during the calibration procedure by techniques well-known in the art, such as high-order statistics. The columns of the matrix **A** consist of the set of linearly independent array response vectors. Note that in the case where the differential delays are small compared with the reciprocal of the signal bandwidth, the invention exploits transmitter movement to decorrelate signals and determine **A**. The matrix **A** is then used to determine a weight matrix **W** such that $\mathbf{WA}=\mathbf{I}$. Using **W** and the sampled signals, one then resolves the q multipath signal parts and calculates their relative time delays, $\tau^{(1)},\ldots,\tau^{(q)}$. These time delays and the matrix **A** associated with each location are stored in the calibration database. (Since the signal subspace may be determined from **A**, it is not necessary to store the signal subspace if **A** is stored.) When attempting to locate a transmitter with an unknown location, this additional information in the database may be used to improve the accuracy as follows. After determining a set of likely locations using subspace matching, the stored matrix **A** is used to calculate differential time delays based on the sampled signals, as was done in the calibration procedure. These time delays are then compared with the stored time delays to help resolve location ambiguities.

In the case of a CDMA implementation, the blind estimation technique is not required to obtain the time delays $\tau^{(1)},\ldots,\tau^{(q)}$ since this delay information is provided by the signal acquisition, as was described above. Preferably, the time delay signature is represented as a vector defined by $\tau=[\tau^{(1)},\ldots,\tau^{(q)}]^T$. In addition, a CDMA implementation also provides the code correlation data represented as a matrix $\mathbf{C}=[\mathbf{C}^{(1)T},\ldots,\mathbf{C}^{(p)T}]^T$, where $\mathbf{C}^{(j)}=[C^{(j)}(\tau_1),\ldots,C^{(j)}(\tau_L)]^T$. The subspaces $\mathbf{U}^{(1)},\ldots,\mathbf{U}^{(q)}$, together with the matrix **C** and the vector τ, are used to construct a signal signature that captures essential location-dependent information about the spatial channel for the mobile at its current location. In addition, the covariance matrices $\mathbb{R}^{(1)},\ldots,\mathbb{R}^{(q)}$ can also be used as part of the signature.

BNSDOCID: <WO_____0034799A1_I_>

The signal signature in the database for each location $\mathbf{d}_k$ preferably includes the calibrated spatial signature $\mathbf{S}_k=\{\mathbf{U}_k^{(1)},...,\mathbf{U}_k^{(q)};\ \mathbf{R}_k^{(1)},...,\mathbf{R}_k^{(q)}\}$, the calibrated temporal delay signature $\tau_k=[\tau_k^{(1)},...,\tau_k^{(q)}]^T$, and the calibrated correlation signature $\mathbf{C}_k=[\mathbf{C}_k^{(1)T},...,\mathbf{C}_k^{(p)T}]^T$. It should be emphasized that the signal signature is not limited to the specific examples given here, but may include other characterizations as well. In addition, it should be emphasized that the signal signature need not contain all the above components. For example, it might contain just a single subspace, $\mathbf{U}_k$.

Velocity Estimation

In accordance with the present invention, the real-time velocity of the phone can be determined by the following method. Because the 1 ms batch sampling time is short compared to the movement of the phone, the batch covariance matrix will have, to a close approximation, rank 1. Thus, denoting the batch covariance matrix by $\mathbf{R}$, we may write

$$\mathbf{R} \approx \mathbf{x}\,\mathbf{x}^H, \tag{1}$$

where $\mathbf{x}\,\mathbf{x}^H$ represents the rank-1 approximation to $\mathbf{R}$, i.e. the vector $\mathbf{x}$ is the dominant eigenvector of $\mathbf{R}$ and $\mathbf{x}^H$ is its Hermetian conjugate.

Now consider a second covariance matrix $\mathbf{R}'$ measured a short time δt after $\mathbf{R}$ is measured (e.g., δt=50 ms), again over the course of a 1 ms time interval. Then write

$$\mathbf{R}' \approx \mathbf{x}'\,(\mathbf{x}')^H, \tag{2}$$

where $\mathbf{x}'\,(\mathbf{x}')^H$ represents the rank-1 approximation to $\mathbf{R}'$.

In a typical multipath environment, the physical displacement of the transmitter is, to a first order approximation, proportional to the angular displacement between the vectors $\mathbf{x}$ and $\mathbf{x}'$ in signal space. The greater the displacement of the phone during the interval δt, the greater the angle between $\mathbf{x}$ and $\mathbf{x}'$. Thus, in order to estimate the velocity of the phone, it would be desirable to measure the angular velocity of the signal vector by comparing two consecutive

BNSDOCID: <WO_____0034799A1_i_>

measurements of $\mathbf{x}$ separated by a known time interval δt. The computation of $\mathbf{x}$, however, would involve a computationally expensive calculation of the eigenvalues and eigenvectors of $\mathbf{R}$. Accordingly, the present inventors have developed the following technique for determining velocity without the need for such complex computations.

Define

$$\delta\mathbf{R} = \mathbf{R}' - \mathbf{R} \tag{3}$$

and

$$\delta\mathbf{x} = \mathbf{x}' - \mathbf{x}. \tag{4}$$

Using these definitions, we may rewrite eq. 2 as

$$\mathbf{R} + \delta\mathbf{R} = (\mathbf{x} + \delta\mathbf{x})(\mathbf{x} + \delta\mathbf{x})^H. \tag{5}$$

Expanding the right-hand side of eq. 5, we obtain

$$\mathbf{R} + \delta\mathbf{R} = \mathbf{x}\,\mathbf{x}^H + \delta\mathbf{x}\,\mathbf{x}^H + \mathbf{x}\,\delta\mathbf{x}^H + \delta\mathbf{x}\,\delta\mathbf{x}^H. \tag{6}$$

Subtracting eq. 1 from this, we then get

$$\delta\mathbf{R} = \delta\mathbf{x}\,\mathbf{x}^H + \mathbf{x}\,\delta\mathbf{x}^H + \delta\mathbf{x}\,\delta\mathbf{x}^H. \tag{7}$$

Now, using eq. 7, one can calculate $\|\delta\mathbf{R}\|^2 = \mathrm{Tr}(\delta\mathbf{R}\,\delta\mathbf{R}^H)$ and $|\mathrm{Tr}(\delta\mathbf{R})|^2$, and then subtract the resulting expressions to yield

$$\|\delta\mathbf{R}\|^2 - |\mathrm{Tr}(\delta\mathbf{R})|^2 = 2|\mathbf{x}|^2|\delta\mathbf{x}|^2 - 2|\mathbf{x}^H\,\delta\mathbf{x}|^2. \tag{8}$$

Define

$$\delta\mathbf{x}_T = \delta\mathbf{x} - (\delta\mathbf{x}^H\,\mathbf{x}_U)\,\mathbf{x}_U, \tag{9}$$

where $\mathbf{x}_U = \mathbf{x}/|\mathbf{x}|$ is the unit vector in the direction of $\mathbf{x}$. Thus, $\delta\mathbf{x}_T$ is $\delta\mathbf{x}$ minus the portion of $\delta\mathbf{x}$ in the direction of $\mathbf{x}$. We can therefore think of $\delta\mathbf{x}_T$ as the tangential component of $\delta\mathbf{x}$.

Using eq. (9) we can rewrite eq. (8) as

BNSDOCID <WO_ _ _ _ 0034799A1_I_>

$$\| \delta \mathbf{R} \|^2 - |\mathrm{Tr}(\delta \mathbf{R})|^2 = 2|\mathbf{x}|^2 |\delta \mathbf{x}_T|^2. \quad (10)$$

Let $\delta\theta$ be the angle between $\mathbf{x}'$ and $\mathbf{x}$. Then, assuming $\delta \mathbf{x}_T << \mathbf{x}$, we have $|\delta \mathbf{x}_T|/|\mathbf{x}| = \tan(\delta\theta) \approx \delta\theta$, and hence $|\delta \mathbf{x}_T| \approx |\mathbf{x}|\, \delta\theta$. Substituting this in eq. 10, we get

$$\| \delta \mathbf{R} \|^2 - |\mathrm{Tr}(\delta \mathbf{R})|^2 \approx 2|\mathbf{x}|^4 (\delta\theta)^2. \quad (11)$$

Finally, because $\mathbf{R} = \mathbf{x}\, \mathbf{x}^H$, we have $\mathrm{Tr}(\mathbf{R}) = \mathrm{Tr}(\mathbf{x}\, \mathbf{x}^H) = |\mathbf{x}|^2$, and hence $|\mathbf{x}|^4 = [\mathrm{Tr}(\mathbf{R})]^2$. Substituting this in eq. 11 gives

$$\| \delta \mathbf{R} \|^2 - |\mathrm{Tr}(\delta \mathbf{R})|^2 \approx 2[\mathrm{Tr}(\mathbf{R})]^2 (\delta\theta)^2. \quad (12)$$

Solving this for $\delta\theta$ we obtain

$$\delta\theta \approx \sqrt{ \{ \| \delta \mathbf{R} \|^2 - |\mathrm{Tr}(\delta \mathbf{R})|^2 \} / 2[\mathrm{Tr}(\mathbf{R})]^2 } \quad (13)$$

Notice that the right-hand side of this equation can be calculated with simple algebraic operations involving only $\delta\mathbf{R}$ and $\mathbf{R}$. This equation, therefore, is a function that gives a simple and computationally efficient way to compute the angular change in the signal vector directly from the measured batch covariance matrices.

The physical displacement of the phone, to first order approximation, is proportional to the angular displacement of the signal vector. To estimate the magnitude of the velocity of the phone at a time t, therefore, the following expression is computed:

$$|v_t| = \alpha\, \delta\theta/\delta t\ , \quad (14)$$

where δt is the time difference between the measurements of $\mathbf{R}'$ and $\mathbf{R}$, and α is a constant of proportionality that is determined empirically. Notice that $|v_t|$ is an estimate of the velocity over the period δt which is typically 5 to 50 ms. To obtain the average velocity over a longer period, say one second, one can use any of the averaging and smoothing techniques described above to a collection of velocity estimates collected within the longer period. It will be appreciated that other equations similar to eqs. 13 and 14 may be used to make similar

BNSDOCID: <WO_____0034799A1_I_>

estimates of the phone velocity. Such substantially similar equations include, for example, mathematically equivalent expressions as well as equations involving a similar combination of the quantities $Tr(\mathbf{R})$, $Tr(\delta\mathbf{R})$, and $Tr(\delta\mathbf{R}\,\delta\mathbf{R}^H)$, or multiples of these quantities. The essential feature of the equation is that it is a simple, closed-form expression in terms of just $\mathbf{R}$ and $\mathbf{R}'$. It will also be appreciated that the direction of the velocity can be estimated by comparing successive locations that have been determined using a location-finding technique. Combining this direction estimate with the magnitude estimate, one may estimate the actual velocity vector, $\mathbf{v}$.

The value of $|v|$ may be used in various ways to improve the performance of a location finding system, as well as for other applications where a knowledge of transmitter movement may be of use.

Calibration Data Collection Procedure

The real-time location determination technique of the present invention makes use of a stored database of calibrated signal signature information and associated locations. The initialization of the system, therefore, preferably involves a calibration procedure in which signal signature and location information is collected. One procedure for collecting this data is as follows.

According to a preferred calibration procedure, a calibration vehicle follows an arbitrary trajectory in the vicinity of a LDA. The vehicle contains an active cellular telephone, a GPS receiver, and a computer for storing the real-time GPS position data of the vehicle along its trajectory. In a typical calibration procedure, the GPS location data along the trajectory are periodically stored on the computer together with the GPS clock time-stamp, typically at intervals on the order of 10 to 100 ms. Meanwhile, the transmitted signal from the cellular phone is received at the LDA, which is also equipped with a GPS receiver, and the signal signature is calculated and stored with the GPS clock time stamp. When the data from the vehicle and the LDA are combined and processed, differential GPS correction data obtained from the LDA are standard techniques are used to determine with high accuracy the vehicle location corresponding to each signature measured at the LDA. The signal signature and location information are then stored as raw calibration data at the base station.

Note that the calibration procedure may include the simultaneous use of several base stations and/or several vehicles moving within the service area. In addition, each vehicle may include several phones operating on different channels. For example, in an implementation for the AMPS standard, a set of 6 phones in a vehicle may be used to transmit at frequency intervals

BNSDOCID <WO____ 0034799A1_I_>

of 2 MHz in order to provide the signature data with frequency coverage across a 10 MHz bandwidth. In this case, the raw calibration data contains 6 signatures corresponding to the 6 frequencies for each location along the vehicle trajectory. It will also be appreciated that the location data may be derived from techniques other than, or in addition to, GPS techniques. For example, the GPS location information can be supplemented by dead reckoning or other navigational information. Such supplementary location information is especially useful in cases where GPS reception is poor. The supplementary location information can be stored with the GPS information and post-processed in an off-line procedure as described above. Alternatively, the location information can be transmitted in real time to the base station during the calibration procedure.

The result of the calibration procedure for a single vehicle with a single phone is a sequential set of raw calibration data for a single trajectory, including a set of covariance matrices $\mathfrak{R}=\{\mathbf{R}_1,...,\mathbf{R}_N\}$ and a set of N corresponding locations $\mathfrak{D}=\{\mathbf{d}_1,...,\mathbf{d}_N\}$. The raw calibration data for a single trajectory is conveniently represented as a single set of ordered pairs, $\mathfrak{T}=\{(\mathbf{R}_1,\mathbf{d}_1),...,(\mathbf{R}_N,\mathbf{d}_N)\}$ indexed by the integers 1 to N. Because locations $\mathbf{d}_i$ and indices i are in one-to-one correspondence, locations may be identified by their corresponding location index, and vice versa. If time delay and/or code correlation information is also measured, this data is also recorded as part of the raw calibration data, so that data for a single trajectory is represented as ordered triplets or ordered quadruplets.

Typically, the calibration procedure will use one or more vehicles following M routes at the same or different times. The raw data collected from M routes is initially stored in a set of M files, which are also represented by $\mathfrak{T}_1,...,\mathfrak{T}_M$. In the case where only covariance matrix information is measured, each file $\mathfrak{T}_m$ contains a temporally sequential list of ordered pairs $(\mathbf{R}_1,\mathbf{d}_1),...,(\mathbf{R}_{N(m)},\mathbf{d}_{N(m)})$, where N(m) is the number of data points in route m. If additional information is measured, each file contains a list of ordered triplets, quadruplets, etc. Each data point, therefore, may be assigned a unique index pair (m,n), where m=1,...,M specifies the route and n=1,...,N(m) specifies the data point within the route, as shown in TABLE 1.

BNSDOCID: <WO_____0034799A1_I_>

TABLE 1

Route 1	Route 2	...	Route m	...	Route M
(1,1)	(2,1)		(m,1)		(M,1)
(1,2)	(2,2)		(m,2)		(M,2)
. :	. :		. :		. :
(1,n)	(2,n)	...	(m,n)	...	(M,n)
. :	. :		. :		. :
(1,N(1))	(2,N(2))		(m,N(m))		(M,N(M))

Note that each route m will usually have a different number of data points N(m). The columns of Table 1, therefore, do not necessarily have the same number of rows.

The raw calibration data is used to form the calibration database which contains K calibrated signal signatures and K associated locations, as shown in TABLE 2. Each of the locations $\mathbf{d}_1,...,\mathbf{d}_K$ may be two or three dimensional, and may be assigned a date upon which it was last calibrated. The signal signature in the database for each location $\mathbf{d}_k$ includes the calibrated spatial subspace information, $\mathbf{S}_k$, and possibly also calibrated temporal delay information, τ_k, and code correlation information, $\mathbf{C}_k$. In a preferred embodiment, the covariance signature is $\mathbf{S}_k=\{\mathbf{U}_k;\ \mathbf{R}_k\}$ or, if multipath parts are resolved, $\mathbf{S}_k=\{\mathbf{U}_k^{(1)},...,\mathbf{U}_k^{(q)};\ \mathbf{R}_k^{(1)},...,\mathbf{R}_k^{(q)}\}$. The temporal delay signature preferably is $\tau_k=[\tau_k^{(1)},...,\tau_k^{(q)}]^T$, and the code correlation signature is $\mathbf{C}_k=[\mathbf{C}_k^{(1)T},...,\mathbf{C}_k^{(p)T}]^T$. It should be emphasized that the signal signature is not limited to the specific examples given here, but may include other characterizations as well.

BNSDOCID: <WO_____0034799A1_I_>

TABLE 2

Location (x, y, z)	Signal Signature		
	Spatial Signature	Temporal Signature	Correlation Signature
$\mathbf{d}_1$	$\mathbf{S}_1$	τ_1	$\mathbf{C}_1$
.	.	.	.
.	.	.	.
.	.	.	.
$\mathbf{d}_K$	$\mathbf{S}_K$	τ_K	$\mathbf{C}_K$

Although the signal signature is a relatively stable function of location, certain changes in the signature may take place which may affect the accuracy of the database of calibrated signal signatures. Such changes may result from the construction or destruction of buildings, or from seasonal changes in the natural environment that affect multipath. Accordingly, the preferred embodiment of the present invention includes techniques for updating the database. A direct method involves performing the calibration procedure at regular intervals or continuously. A cost-effective method of continuous calibration could use a fleet of commercial delivery vehicles. Since such vehicles typically pass through many locations in a localized region, a dedicated calibration system could piggy-back on these vehicles and allow continuous updating of the database. Alternatively, or in addition, specialized vehicles could be sent out periodically to calibrate the service region. It should also be noted that in the case where two or more base stations can simultaneously receive signals from such a calibration unit, its signals can be used by all these base stations to independently calibrate their respective databases.

It is also possible to perform certain database corrections using only signals from regular phones, i.e., without using special calibration vehicles. For example, in the case where seasonal atmospheric changes cause alterations in the multipath, an analysis of signals over the course of several days or weeks can detect a gradual drift in the collective signal signatures. After clearly identifying such a drift, the database signatures can be corrected to compensate for the drift, thus eliminating the need for performing a calibration procedure. A similar technique can be used to correct the database to compensate for signal signature changes caused by the construction or destruction of reflecting objects in the environment. In this case, an analysis of signal signatures over time will display a consistent and gradual shift for those signatures associated with locations in the vicinity of the objects. Having

BNSDOCID: <WO_____0034799A1_I_>

identified such bias, the database can then be compensated accordingly. As an aid to the identification of these environmental biases, base stations may share with each other certain information, e.g. the array response matrix A for a certain location. If both base stations detect a shift, it is likely due to an environmental change. It is a significant feature that high bandwidth communication channels are not required to communicate such information between base stations. In addition, geographical and environmental information can be used to detect these shifts. For example, because major highways do not typically change their location, vehicles being tracked regularly along such a highway will have known trajectories. This trajectory information can be used to identify and correct drifts in the calibration data.

It should be noted that, if a transmitter is stationary and if the differential time delays are much smaller than the reciprocal of the signal bandwidth, then the multipath and direct path signals will be coherent, i.e. they will differ only by a complex scalar factor. In this case, the measured signal subspace will be one-dimensional since the multipath signal does not contribute an additional dimension to the signal subspace. A moving transmitter, on the other hand, will decorrelate the direct and multipath signals over a distance of several wavelengths. For example, two multipath signals from a moving transmitter result in sufficient variation of subsequent signal vectors $\mathbf{x}(t_1),\ldots, \mathbf{x}(t_M)$ to permit the full subspace to be determined. Because it is preferable in the calibration procedure to have signal signatures comprising multi-dimensional subspaces, the calibration procedure of the present invention uses moving vehicles to advantageously decorrelate the direct and multipath signals, thereby permitting determination of the complete signal subspace. Although a stationary transmitter will have only a one-dimensional subspace, because this subspace is within the calibrated signal subspace it is still possible to locate the transmitter, albeit with less certainty than with a moving transmitter. In the case of wideband modulation schemes or frequency hopping, however, the full signal subspace may be determined even for a stationary transmitter.

Calibration Data Processing

In the preferred embodiment, the calibration process includes post-processing to optimize the calibration table. This optimization can involve eliminating redundant information, compressing the data for more efficient storage, and organizing the data for more efficient searching. The raw data contains redundancy due primarily to arbitrary and uncontrolled vehicle speeds and trajectories during the calibration data collection. A stopped or very slow moving vehicle, for example, can result in many redundant data points corresponding to nearly the same location (assuming uniform sampling of data points). Overlapping route trajectories also result in redundant data points. It is therefore advantageous to remove these

BNSDOCID: <WO_____0034799A1_I_>

redundancies in order to generate an optimal calibration table from this raw data. The post-processing method comprises two main parts: data association and grid generation.

DATA ASSOCIATION

The data association procedure is preferably separated into two steps. The first step generates a route neighbor file for each route, while the second step generates a total route neighbor file for all routes.

The first step is performed as follows. For each data point (m,n) in a route m, a list $\mathfrak{N}_m$ of neighboring data points in the same route m is determined. A neighboring data point is defined to be any data point whose location is less than a predetermined distance D from the location of (m,n). In the preferred embodiment, the predetermined distance D is 5 to 10 meters, as this is the approximate accuracy of present GPS receivers and is comparable to a lane width on a typical road. It will be appreciated that other distances D may also be used. In this manner, a route neighborhood file $\mathfrak{N}_m$ for route m is constructed which contains a route neighbor list for each point in the route. TABLE 3 shows an example of a route neighborhood file $\mathfrak{N}_m$ containing route neighbor lists in route m for all points in route m.

TABLE 3: Route Neighbor File

Data Point	Route Neighbor List
(m,1)	(m,2);(m,3)
(m,2)	(m,1);(m,3);(m,4)
⋮	
(m,n)	(m,n-2);(m,n-1);(m,n+1);(m,n+2);(m,n+3)
⋮	
(m,N(m))	(m,N(m)-2);(m,N(m)-1)

Similarly, a route neighborhood file is produced for each of the M routes, resulting in a total of M route neighborhood files $\mathfrak{N}_1,\ldots,\mathfrak{N}_M$. corresponding to each of the routes $\mathfrak{T}_1,\ldots,\mathfrak{T}_M$.

BNSDOCID: <WO_____0034799A1_I_>

After the M route neighborhood files are constructed, the second step of the data association procedure is performed. In this step, a total neighborhood file $\mathfrak{M}$ is constructed, combining information from all the routes. The total neighborhood file $\mathfrak{M}$ contains, for each data point in each route, a total neighbor list. The total neighborhood list for a point contains all the neighbors of the point selected from all data points in all routes. In other words, the total neighbor list for a point contains not only neighbors in the same route, but also neighbors taken from all the other routes. TABLE 4 shows an example of a total neighborhood file.

TABLE 4: Total Neighbor File

Data Point	Total Neighbor List
(1,1)	(1,2);(1,3);(2,12);(2,13);...;(M,54);(M,55)
. :	. :
(1,N(1))	(1,N(1)-1);(2,11);(2,12);...;(M,128)
. . .	. . .
(m,1)	(1,32);(2,78);...;(m,2);(m,3);...;(M,98)
. :	. :
(m,N(m))	(3,45);...;(m,N(m)-1);...;(M-1,32)
. . .	. . .
(M,1)	(1,16);...;(M,2);(M,3)
. :	. :
(M,N(M))	(2,68);(2,69);...;(M,N(M)-2);(M,N(M)-1)

The data processing described above may be performed conveniently through the repeated execution of certain general procedures. A fundamental procedure NEIGHBOR(m_1,n_1,m_2,n_2) checks if data point (m_1,n_1) is a neighbor of (m_2,n_2) by calculating the Euclidean distance between the GPS position $\mathbf{d}_{n_1}$ in route m_1 with GPS position $\mathbf{d}_{n_2}$ in route m_2. If the distance is less than D meters, then the procedure returns a TRUE value; otherwise it returns a false value. Using this procedure, it is straightforward to

BNSDOCID <WO_____0034799A1_I_>

devise a procedure ASSOCIATE_ROUTES(m_1,m_2) that calculates, for each data point in route m_1, the list of neighboring points in route m_2. With m_1=m_2, this procedure can be used to generate the initial route neighborhood files. This procedure can also be used to generate the total neighborhood file from the collection of route neighborhood files by calling the procedure for all cases where $m_1 \neq m_2$ and adding the results of each call to the total neighborhood file. Note that since ASSOCIATE_ROUTES is symmetric, i.e. ASSOCIATE_ROUTES(m_1,m_2) = ASSOCIATE_ROUTES(m_2,m_1), half of the distance calculations need not be performed.

A new route M+1 can be added to the total neighborhood file by calling ASSOCIATE_ROUTES(M+1,m) for all m, and also calling ASSOCIATE_ROUTES(m,M+1) for all m. The first series adds to the new route all neighbors from all existing routes, while the second series of calls adds to the existing routes the neighbors in the new route. Similarly, any route m can be deleted from the total neighborhood file by removing all data points (m,1),...,(m,N(m)) and removing all references to these data points which may be contained in the total neighborhood lists of the data points for the other routes.

GEOGRAPHIC BINNING

The binning procedure uses the contents of the total neighborhood file to construct a collection of geographical bins that represent the area covered by the routes. A bin is defined to be a contiguous geographical region corresponding to a portion of one or more routes. The binning step constructs a collection of bins that correspond to the trajectories of the routes. This collection of bins is a basic component of the final calibration table.

Bins are created sequentially as follows. Initially, all the points in all the routes are tagged as unbinned. A current route variable is initially set to be the first route, and a current point in the current route is initially set to be the initial point in the route. Next, a bin is created that covers a sequence of points starting with the current point and ending with the last successive point in the route that is still within a minimum binning distance from the current point. The bin is assigned a bin location and a signature based on the signatures and locations of the points in the bin, as will be described in more detail below. After the bin is created, the current point is the first point along the route outside the bin. This point becomes the first point for the next bin, and a new bin is created. Bins are successively created along the route in this way until no points remain in the route. After the first route is binned, the current route is then incremented and the binning is repeated in the same manner

BNSDOCID: <WO_____0034799A1_I_>

for the next route. When all routes have been binned in this way, the binning procedure is completed.

The step of creating a single bin is performed as follows. An initial bin point is set to be the current point. All the neighbors of the current point are identified by inspecting the total neighbor list for the point. These neighbors are the points (from any route) that are within a circular neighborhood of the current point. From the list of these neighbors and the current point itself, all the unbinned points are selected. These selected points will be tagged as binned when the bin is completed, so that they are not included as data for other bins. The covariance matrices of these points are then averaged to form an average covariance matrix **R** associated with the current point. From this average covariance matrix, a covariance signature **U** corresponding to this point is then determined. In the preferred embodiment, the covariance signature **U** corresponding to a covariance matrix **R** is the subspace spanned by the dominant eigenvectors of **R**. In addition, the covariance matrix itself can be used as the signature. In the case where there are no unbinned points among the neighbors of the current point, then no signature value is calculated for the point. In the case where the raw data includes time delay and code correlation data, these are also averaged and included in the signature values for the point.

The current point is then incremented to the next point in the current route. If the current point is less than a predetermined binning distance (e.g., 30 meters) from the initial bin point, then the above procedure is repeated to again produce a signature for this point. As a result, one obtains a series of signature values that correspond to points along the current route within the predetermined binning distance of the initial bin point. In general, the binning distance may be selected to be any value greater than the neighbor distance D. In the preferred embodiment, the binning distance is selected to be approximately 50 meters for typical suburban environments. For dense urban environments, the binning distance is preferably 30 meters, while in rural environments the binning distance is preferably 50 meters or more. The signature values in a given bin are then averaged to form a bin signature, and the corresponding locations are averaged to form an average location. A bin location is selected to be the location of the data point closest to the average location.

In a variation of the above step, each of the bins is divided into sub-bins, as illustrated in FIGS. 8A and 8B. The size and shape of the sub-bins is determined by the location information error. In the case of GPS with differential correction, the sub-bins are preferably squares whose size is about 10 m, and the bins have a size of about 30 m. The locations of raw data points are shown in the figures as empty circles. FIG. 8A shows the vehicle routes as dotted lines. FIG. 8B shows detail of a bin containing nine sub-bins. For each sub-bin an average $\underline{\mathbf{R}}$ of the covariance matrices whose locations are within the sub-bin

BNSDOCID: <WO_____0034799A1_I_>

is calculated, together with a centroid $\underline{\mathbf{d}}$ of their locations. These sub-bin centroids are shown in FIG. 8B as solid circles. The same procedure is repeated for the bins. That is, for each bin k an average $\mathbf{R}_k$ of the sub-bin covariances $\underline{\mathbf{R}}$ belonging to the bin is calculated, together with a centroid $\mathbf{d}_k$ of the sub-bin locations $\underline{\mathbf{d}}$ in the bin. The centroid for the bin is shown in the figure as a circled 'x' mark. In addition, a bin signature $\mathbf{U}_k$ is calculated from $\mathbf{R}_k$.

This process is continued until the points in the route are all binned. The other routes are then binned in the same way, resulting in a calibration table that incorporates the data from all the routes. In a preferred embodiment adapted for the AMPS standard, distinct calibration tables are constructed for various different frequency bands in use, e.g., one calibration table for each 3 to 6 MHz band of frequency channels. Thus, the calibration table used to derive a location will be selected in dependence upon the frequency of the signal.

Data from a new route can be incorporated into an existing calibration table without repeating many calculations unnecessarily. In particular, data from a new route is used to create a new total neighbor file in the manner already described. All neighbors of the new route, as well as all the points in the new route, are tagged as unbinned. The calibration table can then be updated by executing the binning procedure for the new route.

The optimized calibration table does not contain temporally or spatially redundant data, and contains accurate data due to averaging within each bin. The method also enjoys the advantage that it does not place constraints upon the specific routes that the vehicles take. Those skilled in the art will recognize that the present method produces a contiguous map of a region, providing accuracies often exceeding those of existing survey maps. Thus, the present method has application to accurate map generation, as well as to location determination techniques.

DATA COMPRESSION AND ORGANIZATION

In order to provide a fast real-time location finding system, it is preferable to efficiently store and organize the calibration data. In addition to the signal signature data, it is also preferable to store some precomputed data that allows for more efficient signature matching during real-time location determination. One example of a type of precomputed data is a set of N-dimensional vectors $\mathbf{w}_1,\ldots,\mathbf{w}_N$, each of which is derived from the covariance matrices $\mathbf{R}_1,\ldots,\mathbf{R}_N$ of all N locations. As will be described in more detail later, these vectors can be used in the signature matching procedure to improve matching accuracy. Each vector is defined by $\mathbf{w}_k=[\mathrm{Tr}\{\mathbf{P}_1\mathbf{R}_k\},\ldots,\mathrm{Tr}\{\mathbf{P}_N\mathbf{R}_k\}]^T$, where $\mathbf{P}_i = \mathbf{U}_i\mathbf{U}_i^H$ is the projector onto the subspace of location i. Thus, each component $\mathrm{Tr}\{\mathbf{P}_i\mathbf{R}_k\}$ of the vector measures the extent to

BNSDOCID: <WO_____0034799A1_I_>

which the covariance matrix $\mathbf{R}_k$ of location k overlaps with the subspace $\mathbf{U}_i$ of location i. This vector can be used to match a real-time covariance matrix $\mathbf{R}$ to a calibrated covariance signature. As will be described in more detail later, a similar vector $\mathbf{w}$ can be calculated from $\mathbf{R}$, and then the quantity $G(\mathbf{w},\mathbf{w}_k) = |\mathbf{w} - \mathbf{w}_k|^2$ is computed to measure the similarity of the real-time signature with calibrated signature k. Because these N-dimensional vectors can take a large amount of storage space, it is desirable to store a more compact representation of their information. It can be shown that $Tr\{\mathbf{P}_i\mathbf{R}_k\} = \mathbf{p}_i^T\mathbf{r}_k$, where $\mathbf{p}_i$ is a p^2-dimensional vector derived from $\mathbf{P}_i$, and $\mathbf{r}_k$ is a p^2-dimensional vector derived from $\mathbf{R}_k$. Moreover, both $\mathbf{p}_i$ and $\mathbf{r}_k$ have real-value components. Thus, $\mathbf{w}_k=[\mathbf{p}_1^T\mathbf{r}_k,...,\mathbf{p}_N^T\mathbf{r}_k,]^T$. More generally, for any pxp covariance matrix $\mathbf{R}$, one can derive a vector $\mathbf{w}=[\mathbf{p}_1^T\mathbf{r},...,\mathbf{p}_N^T\mathbf{r},]^T$, where $\mathbf{r}$ is a p^2-dimensional vector with real-valued components. In order to reduce the required data storage space as well as the time required to search the database during real-time location-finding, the p^2-dimensional vectors $\mathbf{r}_1,...,\mathbf{r}_N$ are stored instead of the N-dimensional vectors $\mathbf{w}_1,...,\mathbf{w}_N$. In addition, the p^2-dimensional vectors $\mathbf{p}_1,...,\mathbf{p}_N$ are also stored. The vectors $\mathbf{w}_1,...,\mathbf{w}_N$ can then be easily reconstructed in real time as needed.

Another similar technique is based on the following method, which is applicable to more general distance functions (i.e., functions that are not necessarily linear). Perform a singular value decomposition (SVD) of the matrix $\mathbf{W} = [\mathbf{w}_1...\mathbf{w}_N]$. Typically, the SVD of this matrix will have an effective rank dramatically smaller than N. For example, typically only about 100 statistically significant dominant eigenvectors result from an SVD of a matrix $\mathbf{W}$ with N=10,000. These dominant eigenvectors span a subspace $\mathbf{W}_S$ of the matrix $\mathbf{W}$. The projector onto this subspace is $\mathbf{P}_{W_S} = \mathbf{W}_S\mathbf{W}_S^H$. Now calculate $\underline{\mathbf{w}} = \mathbf{W}_S^H\mathbf{w}$ and $\underline{\mathbf{w}}_k = \mathbf{W}_S^H\mathbf{w}_k$ for k=1,....N. The vector $\underline{\mathbf{w}}$ is the projection of the N-dimensional vector $\mathbf{w}$ onto the subspace spanned by the dominant eigenvectors of $\mathbf{W}$. Similarly, vector $\underline{\mathbf{w}}_k$ is the projection of the N-dimensional vector $\mathbf{w}_k$ onto the same subspace. The dimension of this subspace is typically about two orders of magnitude smaller than N. Consequently, the vectors $\underline{\mathbf{w}}$ and $\underline{\mathbf{w}}_k$ also have dimension two to three orders of magnitude less than N. Yet, because the other dimensions may be neglected, no significant information is lost by replacing the original N-dimensional vectors by these lower-dimensional vectors. Consequently, the size of the calibration table can be reduced by more than two orders of magnitude by storing these reduced vectors rather than their N-dimensional counterparts.

The matching of the vector $\mathbf{w}$ with vectors in the set of N vectors $\mathbf{w}_1,...,\mathbf{w}_N$ is performed by evaluating $G(\mathbf{w},\mathbf{w}_k)$ for k=1,....N. It can be seen, however, that

BNSDOCID <WO_____0034799A1_I >

$$G(\mathbf{w},\mathbf{w}_k) = \| \mathbf{w} - \mathbf{w}_k \|^2 \approx \| \mathbf{W}_S\mathbf{w} - \mathbf{W}_S\mathbf{w}_k \|^2 = \| \underline{\mathbf{w}} - \underline{\mathbf{w}}_k \|^2.$$

Thus, the compact vectors $\underline{\mathbf{w}}$ and $\underline{\mathbf{w}}_k$ can be used directly in the real time calculation, rather than the original vectors. Because the dimensions of $\underline{\mathbf{w}}$ and $\underline{\mathbf{w}}_k$ are much smaller than the dimensions of $\mathbf{w}$ and $\mathbf{w}_k$, the storage requirements are much smaller and the calculation speeds are much faster.

Estimation of Antenna Array Manifold Data

The set of array manifold vectors $\{\mathbf{a}(\theta) : \theta=\theta_1,...,\theta_L\}$ of antenna arrays are conventionally determined empirically through array calibration procedures well-known in the art. The calibration data of the present invention, however, allows the array manifold vectors to be determined without conventional array calibration procedures. Instead, the array manifold vectors can be calculated from the signature calibration table, as follows. As shown in FIG. 9, the service region of an LDA 190 contains a set of calibrated locations, shown in the figure as circled 'x' marks. Consider an angular sector of the service region having an angular width $\Delta\theta$ centered around an angle θ. In the preferred embodiment the value for $\Delta\theta$ is approximately one degree. This particular value, however, is not necessary for the practice of the present invention, and various other values will work as well. Contained within this sector are several calibrated locations $\mathbf{d}_1,...,\mathbf{d}_K$, having corresponding calibrated signature subspaces $\mathbf{U}_1,...,\mathbf{U}_K$. After these subspaces have been identified for a given sector, each subspace $\mathbf{U}_i$ is examined to determine if its largest eigenvector $\mathbf{e}_i$ has a magnitude significantly larger than the other eigenvectors in the same subspace. For example, if the largest eigenvector exceeds the next largest eigenvector by a predetermined factor, then it is considered to be significantly larger than all the other eigenvectors. Preferably, this predetermined factor is at least 2, but is more preferably 10 or more. The result of this examination of the subspaces results in a set of largest eigenvectors. For example, if all the subspaces have significantly large first eigenvectors, then the set of largest eigenvectors will contain one eigenvector from each subspace, $\mathbf{e}_1,...,\mathbf{e}_K$. These eigenvectors are then combined to form an average eigenvector for the sector, $\mathbf{e}_{av}$. Various techniques can be used to determine $\mathbf{e}_{av}$ from $\mathbf{e}_1,...,\mathbf{e}_K$. For example, the K eigenvectors can be stacked in a matrix, and the singular value decomposition of the matrix can be calculated to determine the best rank-1 approximation to the matrix. This best rank-1 approximation is then used as the value for $\mathbf{e}_{av}$. A more accurate estimate of $\mathbf{e}_{av}$ can be determined by using the initial estimate of $\mathbf{e}_{av}$ to exclude some outlying eigenvectors from the set of K eigenvectors and then repeating the procedure. The final value of $\mathbf{e}_{av}$ for the

BNSDOCID: <WO_____0034799A1_I_>

sector is used as an initial estimate for the array calibration vector $\mathbf{a}(\theta)$ for the sector. Similar estimates are made for all the sectors around the LDA. These estimates then may be improved by imposing a continuity condition on the values of $\mathbf{a}(\theta)$ with respect to θ. The above procedure, therefore, can derive a set of L array calibration vectors $\mathbf{a}(\theta_1),...,\mathbf{a}(\theta_L)$ for a set of L directions $\theta_1,...,\theta_L$ around a LDA.

It should be noted that this technique for generating array calibration vectors from the location database has several important advantages over known methods of generating array calibration vectors. Prior techniques typically require that the array be installed and measured in a special antenna range. The known methods are not useful, therefore, with many existing antenna arrays. The present method, on the other hand, is effective with existing antenna arrays since no restriction is being made on the nature of the array or its supporting structure.

Location Estimation

The approach of the present location determination technique is to measure a real-time signature of a mobile device whose location is unknown, then match the signature with similar signatures in the calibration database. Because the signatures are highly correlated with location, the calibrated locations of the matching signatures will correspond to the location of the mobile device. Various matching procedures can be used that involve different types of signatures, different matching techniques, search optimization algorithms, ambiguity reduction techniques, and so on. The particular combination of techniques used for matching will depend on the type of signature information available, the characteristics of the wireless environment, and other factors.

Covariance Matrix ,Time Delay, and Correlation Matching

In a preferred embodiment, the location of a mobile device in the service area may be determined as follows. Signals originating from the device at an unknown location are received at the LDA, and a signal signature is determined from the signals as described in detail above. In one embodiment of the invention, the signature comprises a covariance signature S that includes a signal covariance matrix $\mathbf{R}$ and a signal subspace $\mathbf{U}$. In another embodiment, the signature includes a signal covariance matrix and signal subspace for each of q multipath parts of the signal, i.e., $\mathbf{S}=\{\mathbf{U}^{(1)},...,\mathbf{U}^{(q)}; \mathbf{R}^{(1)},...,\mathbf{R}^{(q)}\}$. In yet another embodiment, the signature also includes time delay and/or code correlation information. Using the calibration table described above, the signal signature is then compared using a

BNSDOCID: <WO_____0034799A1_I_>

measure of similarity to a set of calibrated signal signatures which correspond to a set of N calibrated locations.

The measure of similarity between the measured spatial signature $\mathbf{S}$ and a given calibrated spatial signature $\mathbf{S}_k$ may be determined by evaluating a function $D(\mathbf{S},\mathbf{S}_k)$ which maps two signatures to a number indicating their degree of closeness to each other. The value of $D(\mathbf{S},\mathbf{S}_k)$ represents the likelihood that a measured covariance signature $\mathbf{S}$ matches a calibrated signature $\mathbf{S}_k$. Because each signature $\mathbf{S}_k$ corresponds to a known location $\mathbf{d}_k$, the value of $D(\mathbf{S},\mathbf{S}_k)$ also represents the likelihood that the transmitter is at location $\mathbf{d}_k$. According to one suitable definition, $D(\mathbf{S},\mathbf{S}_k)$ is the sum of the traces of the projections of the measured finger covariance matrices $\mathbf{R}^{(1)},\ldots,\mathbf{R}^{(q)}$ onto the corresponding calibrated subspaces, which can be calculated as follows:

$$D(\mathbf{S},\mathbf{S}_k) = \sum_{i=1}^{q} \mathrm{Tr}\, \mathbf{P}_k^{(i)}\, \mathbf{R}^{(i)},$$

where $\mathbf{P}_k^{(i)}=\mathbf{U}_k^{(i)}[\mathbf{U}_k^{(i)}]^H$ is the projector onto the subspace $\mathbf{U}_k^{(i)}$. In the case where no multipath parts are resolved (i.e., $q=1$), this formula simplifies to $D(\mathbf{S},\mathbf{S}_k) = \mathrm{Tr}\, \mathbf{P}_k\, \mathbf{R} = \mathbf{p}_k{}^T\mathbf{r}$, where the vectors $\mathbf{p}_k$ and $\mathbf{r}$ are the p^2-dimensional vectors discussed above in relation to the calibration table data compression. Note that this particular measure does not require that the subspace $\mathbf{U}$ be calculated from $\mathbf{R}$, i.e., it does not require calculating the dominant eigenvectors of $\mathbf{R}$. It will be appreciated that $D(\mathbf{S},\mathbf{S}_k)$ can be defined in many other ways to measure how closely the signal covariance resembles any given calibrated signature. For example, $D(\mathbf{S},\mathbf{S}_k)$ can be defined to be the mean of the squares of the cosines of the principal angles between the subspaces $\mathbf{U}$ and $\mathbf{U}_k$.

In addition to the covariance signature, the signal signature may also include a temporal code correlation for the signal, $\mathbf{C}=[\mathbf{C}^{(1)T},\ldots,\mathbf{C}^{(p)T}]^T$, which can be compared with similar code correlations in the calibration database. This comparison can be made using, for example, a least-squares difference measure. For example, the difference measure between the measured correlation signature $\mathbf{C}=[\mathbf{C}^{(1)T},\ldots,\mathbf{C}^{(p)T}]^T$ and a calibrated correlation signature $\mathbf{C}_k=[\mathbf{C}_k{}^{(1)T},\ldots,\mathbf{C}_k{}^{(p)T}]^T$ may be calculated as follows:

$$D'(\mathbf{C},\mathbf{C}_k) = |\mathbf{C}-\mathbf{C}_k|^2 = \sum_{j=1}^{p} |\mathbf{C}^{(j)} - \mathbf{C}_k{}^{(j)}|^2,$$

BNSDOCID: <WO_____0034799A1_I_>

where

$$| \mathbf{C}^{(j)} - \mathbf{C}_k^{(j)} |^2 = \sum_{i=1}^{L} | C^{(j)}(\tau_i) - C_k^{(j)}(\tau_i) |^2.$$

This function $D'(\mathbf{C},\mathbf{C}_k)$ provides a measure of the "distance" between the measured signal code correlation and the code correlation of calibrated location k. It can easily be transformed into a measure of similarity by any of various mappings from $[0,\infty) \rightarrow (0,1]$ that map 0 to 1 (e.g., $f(x)=1/(1+x)$). This separate measure can then be used in addition to $D(\mathbf{S},\mathbf{S}_k)$ in order to improve the accuracy of the match, for example, by keeping only those locations that closely match the measured signal with respect to both code correlation and covariance subspace measures.

Similarly, the temporal delay signature $\tau=[\tau^{(1)},...,\tau^{(q)}]^T$ of the measured signal can be compared to the calibrated temporal delay signature $\tau_k=[\tau_k^{(1)},...,\tau_k^{(q)}]^T$ to provide an additional measure $D''(\tau,\tau_k)$ of difference. For example, one may calculate

$$D''(\tau,\tau_k) = |\tau-\tau_k|^2 = \sum_{i=1}^{q} | \tau^{(i)} - \tau_k^{(i)} |^2.$$

This additional measure of difference (or the corresponding measure of similarity) can then be used in addition to $D(\mathbf{S},\mathbf{S}_k)$ and $D'(\mathbf{C},\mathbf{C}_k)$ in order to further improve the accuracy of the match, for example, by keeping only those locations that closely match the measured signal with respect to all three measures. More generally, it will be appreciated that various combinations of the measures $D(\mathbf{S},\mathbf{S}_k)$, $D'(\mathbf{C},\mathbf{C}_k)$, and $D''(\tau,\tau_k)$ can be used in various ways to select locations appropriately. It will also be appreciated that other formulas may be used to define these functions, provided they measure the degree of similarity or difference between signatures.

In one matching technique of the invention, the measures of similarity and/or difference are used to select a set of most likely locations as follows. First, by evaluating $D(\mathbf{S},\mathbf{S}_k)$ for all N locations, a set of likely locations can be selected by taking, for example, the calibrated locations whose corresponding calibrated spatial signatures are closest to the measured spatial signature. For example, from the locations k=1,....N, select a subset of M locations

BNSDOCID: <WO __ ___ 0034799A1_I_>

$k(1),...,k(M)$ such that $\{D(S,S_{k(1)}),...,D(S,S_{k(M)})\}$ is the set of M largest values in the set $\{D(S,S_1),...,D(S,S_N)\}$. Alternatively, the set of likely locations can be selected by taking those locations whose corresponding calibrated spatial signatures have at least a minimum degree of similarity to the measured signature. For example, select a subset of locations $k(1),k(2),...,k(M)$ such that for $j=1,...M$, $D(S,S_{k(j)}) > h$, where h is a predetermined threshold value. This subset of points is illustrated in FIG. 10 as those points above the dotted line corresponding to the value of h.

Another approach to determining a set of likely locations is to consider the topographical structure of the surface defined by $z = S(x,y)$, where the function S is defined at bin location points $(x_k,y_k)=\mathbf{d}_k$ by $S(x_k,y_k)=D(S,S_k)$, and interpolated at other values of (x,y). The locations corresponding to the minima of this surface S should be the most likely locations of the mobile. A systematic search or other analysis of the surface can be used to produce a small set of locations corresponding to distinct minima of the surface. The best match corresponds to the least of these minima. Most probably, the actual location will have both the least minimum, as well as several other minima in close proximity.

Similar selections of likely locations can be made using the difference measures $D'(C,C_k)$ and $D''(\tau,\tau_k)$, noting that these measures are difference measures, while $D(S,S_k)$ is a similarity measure. The selected sets of likely locations using each measure can then be combined or intersected to obtain a single set of likely locations.

SIGNATURE SPECTRUM MATCHING

Another matching technique can be used in addition to the above techniques to further improve the performance of the location finding system. To describe this technique, we first define an N-dimensional vector **w** by

$$\mathbf{w} = [D(S,S_1),...,D(S,S_N)]^T.$$

The vector **w** characterizes how the spatial signature **S** compares to the entire set of calibrated spatial signatures $S_1,...,S_N$. This vector **w** can be thought of as describing a "spectrum" of the signature **S** with respect to the entire set of calibrated signatures $S_1,...,S_N$. This spectrum can be visualized as the graph of the function $D(S,S_k)$ with respect to k, as illustrated in FIG. 10. Similarly, define a set of N vectors $\mathbf{w}_1,...,\mathbf{w}_N$ by

$$\mathbf{w}_k = [D(S_k,S_1),...,D(S_k,S_N)]^T.$$

BNSDOCID: <WO_____0034799A1_I_>

The vector $\mathbf{w}_k$ characterizes how the calibrated signature S_k compares to the entire set of calibrated spatial signatures $S_1,...,S_N$. In other words, the vector $\mathbf{w}_k$ is a "spectrum" of the calibrated signature S_k with respect to the entire set of calibrated signatures, and corresponds to a graph similar to that shown in FIG. 10 for $\mathbf{w}$.

Because the spectrum $\mathbf{w}$ of a signature contains information about the signature, it can be used for matching. In particular, the spectrum $\mathbf{w}$ of a measured signature can be compared with the spectrum $\mathbf{w}_k$ of each calibrated signature. For example, FIG. 11 shows the graph for the spectrum $\mathbf{w}$ superimposed upon the graphs for spectra $\mathbf{w}_{k(1)}$ and $\mathbf{w}_{k(2)}$. As is evident from the figure, the spectrum $\mathbf{w}_{k(2)}$ is more similar to the spectrum $\mathbf{w}$ than to the spectrum $\mathbf{w}_{k(1)}$. Thus, location k(2) is more likely than location k(1). More generally, by comparing the spectrum $\mathbf{w}$ with each calibrated spectrum $\mathbf{w}_k$, a set of likely locations can be selected, just as the signatures themselves were compared. For example, define a measure of difference G between a real-time signature spectrum $\mathbf{w}$ and a calibrated signature spectrum $\mathbf{w}_k$ using some difference measure. There are many well known techniques for measuring the distance between two vectors. For example, one may define the distance between the vectors by a least-squares difference as follows:

$$G(\mathbf{w},\mathbf{w}_k) = |\mathbf{w} - \mathbf{w}_k|^2 = \sum_{i=1}^{N} |D(S,S_i) - D(S_k,S_i)|^2.$$

A set of likely locations can be selected using this measure of difference alone or in combination with the various measures described above.

The above technique is advantageous and reduces location ambiguity because the comparison between the vectors $\mathbf{w}$ and $\mathbf{w}_k$ includes comparative information from all N components of the vectors, and does not merely select locations based on the components of $\mathbf{w}$ alone. Thus, more information is used to compare the measured signal covariance with the calibrated subspaces, and a more accurate match can be obtained.

Moreover, one can analogously define vectors $\mathbf{w}'$ and $\mathbf{w}''$ from $D'(\mathbf{C},\mathbf{C}_k)$ and $D''(\tau,\tau_k)$, respectively, to obtain functions $G'(\mathbf{w}',\mathbf{w}_k')$ and $G''(\mathbf{w}'',\mathbf{w}_k'')$, respectively. The functions G, G′ and G″ can then be used individually or together to select a set of most likely locations using techniques described above in relation to functions D, D′ and D″. It will be appreciated

BNSDOCID: <WO_____0034799A1_I_>

that various combinations of the measures D, D′, D″, G, G′, and G″ can be used in various ways to select locations using the techniques described above, or similar techniques. It will also be appreciated that other formulas may be used to define these measures.

CALCULATION OPTIMIZATIONS

It should be noted that, in the case where $D(\mathbf{S},\mathbf{S}_k) = Tr\{\mathbf{P}_k\mathbf{R}\}$, this similarity measure can be easily computed by $D(\mathbf{S},\mathbf{S}_k) = \mathbf{p}_k^T\mathbf{r}$, where $\mathbf{p}_k$ is a p^2-dimensional vector derived from $\mathbf{P}_k$, and $\mathbf{r}$ is a p^2-dimensional vector derived from $\mathbf{R}$. Moreover, both $\mathbf{p}_k$ and $\mathbf{r}$ have real-value components. Thus, $\mathbf{w}_k=[\mathbf{p}_1^T\mathbf{r}_k,...,\mathbf{p}_N^T\mathbf{r}_k,]^T$ and $\mathbf{w}=[\mathbf{p}_1^T\mathbf{r},...,\mathbf{p}_N^T\mathbf{r},]^T$. This technique simplifies the computation of $D(\mathbf{S},\mathbf{S}_k)$ and $G(\mathbf{w},\mathbf{w}_k)$, as well as reducing the amount of storage space required for precomputed values.

With these simplified representations of $\mathbf{w}$ and $\mathbf{w}_k$, the closeness of the measured signature and a calibrated signature can be calculated, as before, with the function $D(\mathbf{w},\mathbf{w}_k)$. This measure, however, was defined using all N components of the vectors. It is desirable, therefore, to simplify the calculation of D to use the simplified representations of the vectors. One such technique for simplifying the calculation of D is as follows. Define the matrices $\underline{\mathbf{P}}^T = [\mathbf{p}_1^T...\mathbf{p}_N^T]$ and $\underline{\mathbf{R}} = [\mathbf{r}_1...\mathbf{r}_N]$. It can then be shown that the N×N calibration table matrix $\mathbf{W} = [\mathbf{w}_1...\mathbf{w}_N]$ is given by $\mathbf{V} = \underline{\mathbf{P}}\ \underline{\mathbf{R}}$, and that

$$\begin{aligned} G(\mathbf{w},\mathbf{w}_k) &= |\mathbf{w} - \mathbf{w}_k|^2 \\ &= |\underline{\mathbf{P}}\mathbf{r} - \underline{\mathbf{P}}\mathbf{r}_k|^2 \\ &= |\mathbf{Q}\mathbf{r} - \mathbf{Q}\mathbf{r}_k|^2, \end{aligned}$$

where $\mathbf{Q}^T\mathbf{Q}$ is the Cholesky factorization of $\underline{\mathbf{P}}^T\underline{\mathbf{P}}$. Since $\mathbf{Q}$ is a $p^2\times p^2$ square matrix, and since both $\mathbf{r}_k$ and $\mathbf{r}$ are p^2-dimensional vectors, it is very simple to store $\mathbf{Q}$ and $\mathbf{Q}\mathbf{r}_k$, and to compute $G(\mathbf{w},\mathbf{w}_k) = |\mathbf{Q}\mathbf{r} - \mathbf{Q}\mathbf{r}_k|^2$. In a typical embodiment, with p=6 and N=10,000, this technique provides more than two orders of magnitude savings in storage and computation.

In an alternate technique, a vector $\underline{\mathbf{w}} = \mathbf{W}_S^H\mathbf{w}$ is the projection of the N-dimensional vector $\mathbf{w}$ onto the subspace $\mathbf{W}_S$ spanned by the dominant eigenvectors of $\mathbf{W} = [\mathbf{w}_1...\mathbf{w}_N]$. Similarly, vector $\underline{\mathbf{w}}_k = \mathbf{W}_S^H\mathbf{w}_k$ is the projection of the N-dimensional vector $\mathbf{w}_k$ onto the same subspace. These vectors have dimensions significantly smaller than N, and thus

BNSDOCID: <WO_____0034799A1_I_>

require less storage space. The matching of the vector w with vectors in the set of N vectors $w_1,...,w_N$ is performed by evaluating $G(w,w_k)$ for k=1,...,N. It can be seen, however, that

$$G(w,w_k) = | w - w_k |^2 \approx | W_S w - W_S w_k |^2 = | \underline{w} - \underline{w}_k |^2.$$

Thus, the computation of $G(w,w_k)$ is significantly simplified.

ANGULAR ENERGY DISTRIBUTION MATCHING

Another approach to signature matching that can be used separately or together with the above techniques involves the use of array manifold vectors $\{a(\theta)\}$. These vectors can be determined using signature calibration information, as described above, or by other known techniques. Once these vectors are known, they can be used for locating a mobile device, as follows. The covariance matrix **R** for the device is measured as described above, and an angular energy distribution $B(\theta)$ for the signal is calculated for each value of θ. Typically, there are approximately 360 values. In the preferred embodiment, the angular energy distribution is calculated with a beamforming metric, as follows:

$$B(\theta) = a(\theta)^H \mathbf{R}\, a(\theta).$$

Alternatively, any high resolution angular energy distribution metric can be used, such as the MUSIC metric,

$$B(\theta) = a(\theta)^H \mathbf{U}\mathbf{U}^H a(\theta),$$

where **U** is the matrix of the dominant eigenvectors of **R**. A similar function is precomputed in advance for each of the calibrated signal covariance matrices $\mathbf{R}_1,...,\mathbf{R}_N$. One then obtains a set of N angular energy distribution functions $B_1(\theta),...,B_N(\theta)$ corresponding to the set of N calibrated locations. For the beamforming metric, one calculates

$$B_k(\theta) = a(\theta)^H \mathbf{R}_k\, a(\theta).$$

Alternatively, for the MUSIC metric one calculates

BNSDOCID: <WO_____0034799A1_I_>

$$B_k(\theta) = \mathbf{a}(\theta)^H \mathbf{U}_k \mathbf{U}_k^H \mathbf{a}(\theta),$$

where $\mathbf{U}_k$ is the matrix of the dominant eigenvectors of $\mathbf{R}_k$. This set of functions $B_1(\theta),...,B_N(\theta)$ is preferably calculated in advance and stored in the calibration database.

In a preferred embodiment, a set of likely locations is determined by comparing the real-time angular energy distribution $B(\theta)$ of the signal from the mobile device with the calibrated angular energy distributions $B_1(\theta)$, ..., $B_N(\theta)$. In the preferred embodiment, these distributions are compared using a statistical measure $\underline{G}$ to measure the similarity between B and each of the calibrated distributions $B_1,....,B_N$. For example, the measure $\underline{G}$ can be defined by the squared difference:

$$\underline{G}(B, B_k) = \sum_\theta | B(\theta) - B_k(\theta) |^2,$$

where the sum is over all angles $\theta=1,...,360$. Alternatively, the metric $\underline{G}$ can be defined as the Kullback-Liebler distance or other metric. Using this similarity measure $\underline{G}$, the set of likely locations is then selected using techniques such as those already described above for the similarity measures D, D′, D″, G, G′, and G″.

ADDITIONAL AMBIGUITY REDUCTION TECHNIQUES

In addition to the above techniques which use various types of signature analysis to resolve ambiguities and improve matching, various other techniques can be used for resolving ambiguities and improve the accuracy of the location determination technique.

Map Comparisons

One technique for reducing ambiguities is to compare the computed likely locations with locations on a map. For example, if the mobile device is known to be moving, it can be assumed that it is most likely in a vehicle constrained to known roads. Thus, any off-road locations can be eliminated or assigned lower likelihoods, especially if the mobile has a large velocity.

Tracking

BNSDOCID: <WO_____0034799A1_!_>

Another technique for resolving ambiguity is a motion tracking system that associates sets of likely locations measured over the course of several seconds or minutes. Using realistic assumptions about the possible changes in phone location over the course of time, the ambiguities in location can be reduced or resolved. For example, it is very unlikely that a phone that has been located in one localized region of the service area will suddenly move to another location several miles away. Motion continuity constraints, in other words, can be used to eliminate or reduce the likelihood of some locations.

A preferred tracking technique will now be described in relation with FIG. 12. The technique includes constructing trajectories, called tracks, from the sets of likely locations over the course of time. These tracks are illustrated in the figure by solid curves. The technique also includes updating the tracks when each consecutive set of likely locations is obtained, and determining the likelihoods of the various tracks. Similar techniques are well-known in the art of military target tracking. The current tracks at any given point in time have an end point corresponding to the most recent location of the track, and a predicted next point. The end points are shown in the figure as solid circles at the right end of the tracks, while the predicted next points are shown as empty circles. Predicted paths, shown as broken lines, connect the end points with the predicted next points. A set of points corresponding to the next measured set of likely locations is then obtained using the matching techniques described above. These measured points are shown as disconnected solid circles. Typically, in the next measured set of likely locations, some points will be close to the predicted next points of the tracks and some others will not be close. Clearly, the points that are close to present tracks are more likely to correspond to the actual location.

Given a set of tracks and a set of likely locations, calculate for each track a predicted next location and compute its variance (e.g., using the well-known Kalman filter). Using the likelihoods of the data points in the track, calculate the a priori probability of the track. Next, among all the possible pairwise associations between predicted next locations of tracks and the measured set of likely locations, determine the association that maximizes the a posteriori probabilities of the pairwise associations, i.e. find the best match between the predicted track locations and the actual locations. Eliminate associations whose distance is greater than a predetermined association threshold. Terminate any tracks that have a predetermined number of consecutive association failures. For any unassociated locations, initiate new tracks. New tracks are considered tentative until they have been confirmed. A tentative track is confirmed when I locations are associated to it out of J consecutive attempts. The integers I and J are design parameters that are selected in dependence upon the particular desired system tracking performance. All the tracks are then filtered using the well-known Kalman filter. The most likely location of the mobile device is the new (filtered) endpoint of the most likely track. Typically, there will be various short tracks corresponding to erroneous

BNSDOCID <WO_____0034799A1 I. >

locations, and one longer track corresponding to the correct location. Other tracking methods may be used to achieve similar results. In addition, this tracking method may be used alone or in conjunction with other methods of reducing ambiguity.

Multi-Site Data

Another technique for resolving ambiguity is to combine sets of likely locations obtained from multiple LDAs. This technique makes use of a set of several LDAs in communication with each other, e.g., through a central hub. As illustrated in FIG. 13, a hub 192 collects location information over communication lines 194 connected to multiple LDAs 196, 198, and 200. Each LDA receives signals from a mobile 202, estimates its current location, and transmits these location estimates over conventional communication lines 194 to the hub 192. Thus, location estimates are derived independently by the multiple LDAs and sent to the hub. In the case where multiple LDAs have determined sets of likely locations for the same mobile device, the hub can intersect or otherwise combine these sets to improve the accuracy of and confidence in the location estimates.

The combination of the sets of likely locations from the distinct LDAs reduces the ambiguity because the propagation paths of the signal to the different LDAs are different and, consequently, only the true location is likely to be common to all sets. As a result, this technique proves highly effective in reducing location ambiguity. It is significant to note that no temporal synchronization between LDAs is necessary because the location information is determined independently at each of the separate LDAs. Conventional location determination techniques, in contrast, often require precise synchronization at multiple base stations in order to measure differential times of arrival between signals received at the different base stations. Such techniques also require high bandwidth communication to send batches of raw timing data to a central hub for location determination computations. The present technique, in contrast, computes the locations from the raw data at each LDA and only sends a small set of likely locations to the hub. Thus, no high speed or high bandwidth communication lines are required.

In an alternate version of multi-site ambiguity reduction, rather than sending likely location data from the LDAs to the hub, signal covariance matrices are sent to the hub and signature matching is performed by the hub rather than by each LDA. In this method, the calibration data for each LDA is stored at the hub rather than at each LDA. This system architecture reduces the hardware cost of the LDAs. The hub, on the other hand, is preferably provided with one or more very powerful computers, which it uses to perform signature matching for all the LDAs it serves. One approach to performing this matching simply uses the matching process that would have been used by each LDA, then combines the independently derived likely locations to reduce ambiguity, just as described above. Alternatively, the hub can take

BNSDOCID: <WO_____0034799A1_I_>

advantage of the fact that it has signature data from multiple LDAs to perform more sophisticated matching. One technique for implementing this idea is as follows. Let $G_1(k)$ be a difference measure between a signature measured at a first LDA and the calibrated signature corresponding to location k for the same LDA. For example, if $\mathbf{w}$ and $\mathbf{w}_k$ are the measured and calibrated spectra at this LDA, then $G_1(k)$ might be defined by $G_1(k) = G(\mathbf{w},\mathbf{w}_k)$. One can also define $G_1(k)$ using various other measures described above. Similarly, define $G_2(k)$ to be a difference measure between a signature measured at a second LDA and the calibrated signature corresponding to the same location k for this second LDA. Now, rather than independently selecting sets of likely locations using $G_1(k)$ and $G_2(k)$ separately, one can select a set of likely locations in one step using $G_1(k)$ and $G_2(k)$ together. For example, select those locations k such that $G_1(k) + G_2(k) < F$, where F is a predetermined difference threshold. Alternatively, one could select the set of M locations $k(1),...,k(M)$ having the smallest sums $G_1(k) + G_2(k)$ of all the locations. Because this technique simultaneously compares the signals with both the calibrated databases, it provides a higher degree of accuracy and ambiguity reduction than the technique of comparing signals with the databases independently. It will be appreciated by those skilled in the art that this technique generalizes in a straightforward way to three or more base stations. It will also be appreciated that the transmission of covariance matrices from the LDAs to the hub does not require large bandwidth since each covariance matrix is typically only a six-by-six matrix, and is typically sent once every 5 seconds.

EFFICIENT SEARCHING TECHNIQUES

In order to improve the computational efficiency of the real-time signature matching, the calibration data can be organized, and a searching technique can be used to efficiently search the organized calibration information. According to one such technique, a hierarchical tree structure, as shown in FIG. 14, is associated with the set of N calibrated locations. The tree structure is composed of a number of nodes, shown as circles in the figure. Each node is connected to one higher level node (except for the top node, shown containing the numeral "1".) and each node is also connected to a number of lower level nodes (except the bottom nodes, shown containing the numerals "1/1024"). For simplicity of illustration, the tree in the figure has only two branches descending from each node. Preferably, there are actually 4 to 16 branches descending from each node, and 3 to 6 levels of nodes. The leaf nodes of the tree at the bottom level have no lower level nodes.

Using well-known data clustering methods, each location k is assigned to at least one leaf node at the lowest level of the tree. Typically, each leaf node contains 10 to 30 locations. A node at the second-lowest level contains the union of the points contained in the lower-level

BNSDOCID: <WO 0034799A1_I >

nodes connected to it. A similar relationship holds between other nodes in the tree. Thus, the highest node contains the indices of all N locations in the calibration table, while the second-highest nodes contain (not-necessarily disjoint) subsets of these N indices. As discussed earlier, signatures preferably are matched by comparing the spectrum $\mathbf{w}$ of a measured signature with the spectrum $\mathbf{w}_k$ of calibrated signatures. This tree-searching technique, however, can be used with other matching techniques as well. The following description of the tree search, however, assumes for the sake of illustration that the matching is performed using the signal spectrum.

Each node q of the tree has associated with it a mean spectrum vector $\mathbf{w}_{av}(q)$ calculated by averaging all the vectors $\mathbf{w}_k$ corresponding to locations in the node q. Also associated with the node q are precomputed node parameters

$$G_{max}(q) = \max \{ G(\mathbf{w}_{av}(q),\mathbf{w}_k) \text{ such that location k is in node q} \},$$
$$G_{min}(q) = \min \{ G(\mathbf{w}_{av}(q),\mathbf{w}_k) \text{ such that location k is in node q} \}.$$

Each leaf node also has associated with it a set of precomputed distances $G(\mathbf{w}_i,\mathbf{w}_j)$, for all pairs $\mathbf{w}_i,\mathbf{w}_j$ in the node such that $i > j$. These precomputed values are preferably stored with the calibration table data and updated whenever the table is recalibrated.

Using this hierarchical tree-structure and associated data, the process of searching the calibration database for a matching signature is made much more efficient. The preferred searching technique has two phases. The tree search procedure used in the first phase of the matching is an adaptation of the well-known branch and bound algorithm, which traverses the tree and efficiently eliminates all nodes from the search that cannot contain the best match. The elimination of nodes from the search is guided by the following rule. A node q is eliminated if either

$$G(\mathbf{w}, \mathbf{w}_{av}(q)) - G_{max}(q) > B$$

or if

$$G_{min}(q) - G(\mathbf{w}, \mathbf{w}_{av}(q)) > B$$

where B is the distance from $\mathbf{w}$ to the best matching vector among the vectors considered thus far in the tree.

Once a leaf node is reached, an efficient search technique is used to perform the matching within the node. In this technique, the precomputed values $G(\mathbf{w}_i,\mathbf{w}_j)$ for the node are used to quickly eliminate vectors within the node. Note that the storage requirement for the

BNSDOCID: <WO_____0034799A1_I_>

precomputed distances $G(w_i,w_j)$ is quadratic in the number of points in the leaf node. Consequently, the number of points in the leaf nodes is preferably kept small in order to minimize storage requirements. In the preferred embodiment, the number of points in a leaf node is typically between 10 and 30. The elimination of points from the search process within a leaf node q is performed using the following rule. A location $\mathbf{d}_k$ is eliminated if

$$| G(\mathbf{w}, \mathbf{w}_{av}(q)) - G(\mathbf{w}_k, \mathbf{w}_{av}(q)) | > B,$$

or if

$$| G(\mathbf{w}_s,\mathbf{w}_k)- G(\mathbf{w}_s,\mathbf{w}) | > B,$$

where B is the distance from $\mathbf{w}$ to the best match among the set of vectors considered so far in the tree, and where s indexes a special point $\mathbf{d}_s$ selected within the leaf node. Note that this rule does not require any new distance computations beyond the single distance $G(\mathbf{w}_s,\mathbf{w})$. Thus, to quickly eliminate as many points as possible, it is important to choose $\mathbf{d}_s$ so that the distance between w_s and $\mathbf{w}$ is small. In the preferred embodiment, the index s for $\mathbf{d}_s$ is determined by the following approximation criterion:

$$s = \arg \min_s \max_k \{ | G(\mathbf{w}_s,\mathbf{w}_k) - G(\mathbf{w}_s,\mathbf{w}) |^2 \},$$

or

$$s = \arg \min_s \{ \sum_k | G(\mathbf{w}_s,\mathbf{w}_k)- G(\mathbf{w}_s,\mathbf{w}) |^2 \},$$

where the $\max_k$ and the $\sum_k$ are evaluated over all indices k corresponding to points considered so far in the node, and the $\arg \min_s$ function ranges over all indices s corresponding to points not considered yet in the node. This process of point elimination is continued until all the points in the leaf have been either eliminated or considered.

It should be noted that this efficient matching procedure can be easily extended to search for a multiple number of best locations. Alternatively, it can be extended to search for a set of likely locations that satisfy various other criteria. For example, one criterion for a likely location $\mathbf{d}_k$ is that $G(\mathbf{w}_k,\mathbf{w}) < \alpha B$, where B is the distance to the best match and α is a constant greater than 1.

BNSDOCID: <WO_ _ _____0034799A1_I_>

Applications of Location Information

The location determination techniques of the present invention provide real-time location information of mobile devices with accuracy and confidence higher than any similar land-based systems. In addition, in contrast with other systems, the present techniques do not require multiple base stations to determine a location of a mobile device, nor do the present techniques require modification of existing communication system base stations or mobile devices. The present techniques also have the unique advantage that they provide accurate location information in severe multipath environments. Moreover, the accuracy of the present system actually improves in the presence of severe multipath. These and other advantages of the present invention provide a unique and practical solution to the long-standing problem of location determination in wireless communication systems. For the first time, accurate location information can be determined in severe multipath environments and in environments where only one base station can receive a mobile device signal. In addition, this accurate location information can be provided without expensive and elaborate modification or replacement of the existing communications network hardware. This significant improvement in location determination, therefore, provides for the first time techniques that enable the widespread practical realization of many location-based services that benefit the mobile device users, the wireless communications provider, and third parties.

Many useful applications of real time location information have long been known, but have not been commercially realized due to the technical and economic difficulties in obtaining accurate location information. Because the present invention overcomes these difficulties, these desirable applications can now be realized. Although these applications have been described in the prior art in relation to other location determination techniques, to the extent that they are independent of these particular techniques, they may be adapted to make use of the location determination technique of the present invention. In other words, the substitution of the present location finding technique in the place of a prior location finding technique in previously known applications is in most cases an obvious and straightforward modification. It should be noted, however, that because the present technique is based upon a unique approach to the location finding problem, new and useful applications of location information may be conceived that have not been previously known at all.

Wireless Subscriber Services

The accurate and reliable location information provided by the present invention can be used to provide many important services to the mobile device users.

Emergency Dispatching

BNSDOCID: <WO_____0034799A1_I_>

In the event that a wireless device user places a call (e.g., a 911 call) to report an emergency situation, the location of the user can be automatically determined and quickly transmitted to the appropriate authorities to ensure a fast response. In particular, the location information derived from one or more LDAs or from an LDA hub is stored in a real-time location database that is accessible to an emergency dispatcher, e.g., via a computer network connection. Because the caller's location can be tracked in real-time, dispatchers and dispatched emergency vehicles can monitor the location of a user who is not stationary.

Location-Based Service Information

Location information also can be used for location-enhanced information services, e.g., for providing navigational directions and/or traffic condition information to a motorist, or to provide information to a caller about nearby restaurants, gas stations, automatic teller machines, parking lots, movie theatres, and so on. This application can be implemented, for example, by allowing an information vendor to access the real-time location database of callers. The caller location information can then be used to access a location-indexed database of services. The information vendor may be an automated system, e.g., a website or audio information system.

Vehicle Fleet Management

The locations of commercial vehicle fleets can be tracked in real time in order to assist in the management of the vehicles. For example, automatic smart vehicular scheduling can use real-time location information to optimize the use of vehicles based on vehicle locations and the locations of pick-ups or deliveries. The location information can also be used to optimize the routes taken by individual vehicles based on current traffic conditions, road condition, and road construction. Over the long term, location information can also be used to monitor vehicle use and schedule vehicle maintenance.

Tracking People and Objects

Location information also can be used for remotely monitoring the location of children, patients, or other individuals. The person need only wear a small micro-transmitter "tag" that periodically transmits a sufficiently powerful signal to be received by an LDA. A similar "tag" can be attached to objects such as automobiles to assist in locating them in the event that they are lost or stolen. In one possible implementation of this application, the location of the "tag" can be accessed by an authorized subscriber to the service via a website or by telephone.

BNSDOCID: <WO_____0034799A1_I_>

WIRELESS PROVIDER SERVICES

A wireless communication provider can make use of location information to manage wireless networks individually and collectively. For example, location information can be used to assist in hand-off between cells, for automatic trunking, for channel borrowing, for channel capacity optimization, to more efficiently allocate cellular frequencies, and to increase the capacity of the network. The location information can also be used to actively monitor and/or manage the wireless traffic in a network, e.g., to help determine locations where calls are dropped, and to help determine where new base stations are needed. Location information also can be used for location-based billing of cellular calls and to help combat cellular fraud. It also can be used in conjunction with antenna selection techniques to improve individual cell efficiency and capacity. Since antennas at existing cellular base stations have differing directional responses, this technique allows more efficient assignment and use of frequency channels within presently existing cells. In particular, by using location information to pool channels among different sectors, the base station can reduce blocking and increase trunking efficiency.

In addition, location information provided by the present method can be used to improve the quality of and extend the range of existing base stations by beamforming in both receiving and transmission. For receiving, one can use the matrix **A** stored in the calibration table to compute the optimal beamformer. For transmission, one has to compute first the corresponding **A** matrix for the transmission frequency. This can be done as part of the calibration process by probing techniques, i.e., by using several different weight vectors and computing the received signal at the mobile. The matrix **A** can then be calculated from this data off-line.

Also, the location finding method can be used to implement location division multiple access (LDMA) to increase the capacity of cellular systems by allowing several phones to use the same frequency in the same base station service area. In contrast to switched beam techniques which divide the service area into large spatial sectors, LDMA divides the service area into small localized subcells. The subcells are based on the computation of signal signatures, such as the signal subspace. The shapes and sizes of the localized subcells are determined by the multipath reflectors in the environment. In particular, each subcell is characterized by a unique signal signature, so that transmitters in different subcells display distinct signal signatures. Accordingly, the present technique of LDMA often permits the isolation of interferers which might otherwise be in the same switched beam sector.

The present invention has an important advantage over spatial division multiple access (SDMA) techniques in that location-based computations can be done off-line once the

BNSDOCID: <WO_____0034799A1_I_>

calibration database for the base station has been determined. Frequency assignment and beamforming can then be performed quickly in real time since a large amount of preprocessing has been performed in advance. Conventional SDMA techniques, in contrast, require complex computations to be performed in real time.

To implement this scheme, the beamforming in both the reception and transmission is computed from the whole set of **A** matrices corresponding to all the phones using the frequency. To compute the locations of all the users, the preferred embodiment is modified to incorporate techniques, such as those based on high order statistics, that can blindly estimate the array vectors in the presence of statistically independent interferers. This step should be carried out on every batch of 100 samples spanning a 1 ms sampling interval. Beamforming techniques such as this will improve signal to noise ratios, increase the overall capacity of the cellular network, and allow a higher density of base stations to be deployed. Network management techniques are used to coordinate nearby base stations to efficiently share frequencies based on the locations of a set of transmitters.

The foregoing system also provides a means for trouble shooting a base station. The array response matrix **A** can be used to continually monitor the performance of the array of antennas and/or equipment at the base station. For example, sudden and dramatic changes in the elements of the **A** matrix for all locations are likely due to hardware damage to the base station. An examination of the **A** matrices can be used to isolate and identify the damage. Smaller changes in **A** may indicate minor hardware changes in the base station and/or changes in the propagation environment. A comparison of such changes in the array response at one base station with those at another can be used to determine whether or not the changes are isolated to one station or common to all stations in the shared environment.

THIRD-PARTY SERVICES

The location information provided by the present invention can be used by various third-party entities in advantageous ways. For example, highway safety authorities can monitor metropolitan freeway traffic flow in real time, and obtain location-specific information about current traffic flow, congestion, likely accidents, etc. This information can be used to automatically display messages to motorists advising them to take alternate routes to avoid troubled areas, and/or to automatically dispatch police or fire vehicles to investigate likely traffic accidents. This traffic information can also be sent or made available directly to motorists whose locations indicate that they may be affected by the conditions.

BNSDOCID: <WO_____0034799A1_I_>

CLAIMS

What is claimed is:

1. A method for location estimation in a wireless communication system comprising a mobile device and a location determination apparatus, the method comprising:

a) coherently receiving at the location determination apparatus a set of p signals using p receivers coupled to an antenna array comprising p antenna elements;

b) calculating at the location determination apparatus a signal signature for the mobile device from the coherently received set of p signals, wherein the signal signature characterizes a spatial channel between the mobile device and the location determination apparatus; and

c) selecting a set of likely locations from a predetermined set of calibrated locations corresponding to a predetermined set of calibrated signal signatures, wherein selecting the likely locations comprises matching the signal signature with similar calibrated signal signatures.

2. The method of claim 1 wherein the signal signature comprises multipath information of the spatial channel.

3. The method of claim 1 wherein calculating the signal signature comprises collecting a batch of p-dimensional signal vectors derived from the received signals, wherein the p-dimensional signal vectors represent signals originating at the mobile device.

4. The method of claim 3 wherein calculating the signal signature further comprises calculating a signal covariance matrix $\mathbf{R}$, wherein the calculating comprises combining outer products of the p-dimensional signal vectors.

5. The method of claim 1 wherein calculating the signal signature comprises combining signal covariance matrices derived from outer products of the p-dimensional signal vectors.

6. The method of claim 5 wherein combining the signal covariance matrices comprises forming a linear combination of signal covariance matrices weighted by factors derived from an estimated velocity of the mobile device.

BNSDOCID: <WO_____0034799A1_I_>

7. The method of claim 5 wherein the combining comprises estimating a velocity of the mobile device from a pair of recently measured covariance matrices.

8. The method of claim 1 wherein calculating the signal signature comprises comprises resolving each of the p signals into q multipath parts, and calculating q signal covariance matrices for the q multipath parts.

9. The method of claim 1 wherein calculating the signal signature comprises resolving each of the p signals into q multipath parts, estimating q relative temporal delay times of the q multipath parts, and forming a temporal delay signature from the q relative temporal delay times.

10. The method of claim 1 wherein calculating the signal signature comprises deriving p code correlations from the p signals and forming a code correlation signature from the p code correlations.

11. The method of claim 1 wherein the selecting comprises evaluating a function measuring the similarity between the signal signature and a calibrated signal signature.

12. The method of claim 11 wherein evaluating the function comprises computing a trace of a projection of a signal covariance matrix onto a calibrated covariance matrix subspace.

13. The method of claim 1 wherein the calibrated signal signatures comprise a signature chosen from the group consisting of a covariance matrix signature, a temporal delay signature, a code correlation signature, and an angular energy distribution signature.

14. The method of claim 1 wherein the selecting comprises calculating signature similarity functions for multiple signature types and combining the calculated functions.

15. The method of claim 1 wherein the calibrated signal signatures comprise a set of calibrated vectors $\mathbf{w}_k$, wherein each calibrated vector $\mathbf{w}_k$ measures a degree of similarity between a calibrated signal covariance matrix $\mathbf{R}_k$ and a collection of calibrated signal covariance matrices, and wherein the signal signature comprises a vector $\mathbf{w}$ measuring degree of similarity between a signal covariance matrix and the collection of calibrated signal covariance matrices.

16. The method of claim 15 wherein selecting comprises calculating the difference between the vector $\mathbf{w}$ and a calibrated vectors $\mathbf{w}_k$.

BNSDOCID: <WO 0034799A1_I_>

17. The method of claim 15 wherein the dimensions of $\mathbf{w}$ and $\mathbf{w}_k$ are both equal to p^2.

18. The method of claim 1 wherein the selecting comprises searching nodes of a tree-structure indexing the calibrated signatures.

19. The method of claim 18 wherein the searching comprises eliminating nodes of the tree-structure, and searching calibrated locations within remaining nodes.

20. The method of claim 1 wherein the selecting is performed by the location determination apparatus.

21. The method of claim 1 further comprising communicating the calculated signal signature to a hub, and wherein the selecting is performed at the hub, preferably using supplementary signal signatures communicated from additional location finding apparatuses.

22. The method of claim 1 further comprising modifying the set of likely locations based on a comparison of the likely locations with geographical information.

23. The method of claim 1 further comprising tracking subsequent sets of likely locations.

24. The method of claim 1 further comprising updating the predetermined set of calibrated signal signatures based on the signal signature and known geographic information.

BNSDOCID: <WO_____0034799A1_I_>

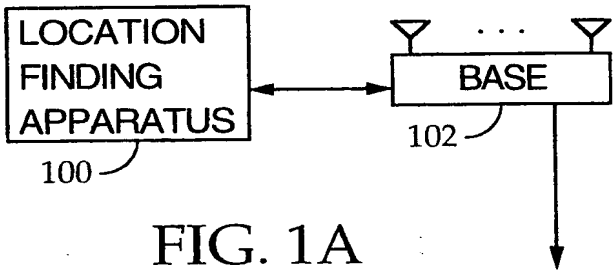


FIG. 1A

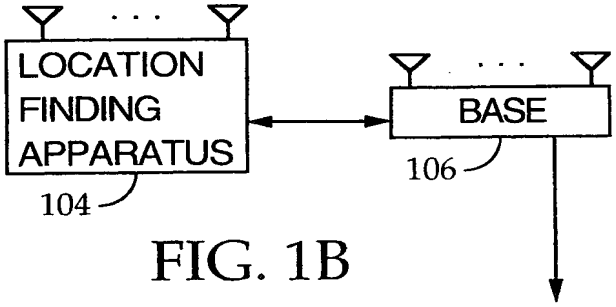


FIG. 1B

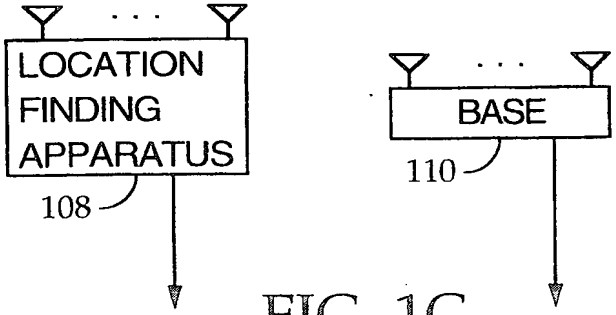


FIG. 1C

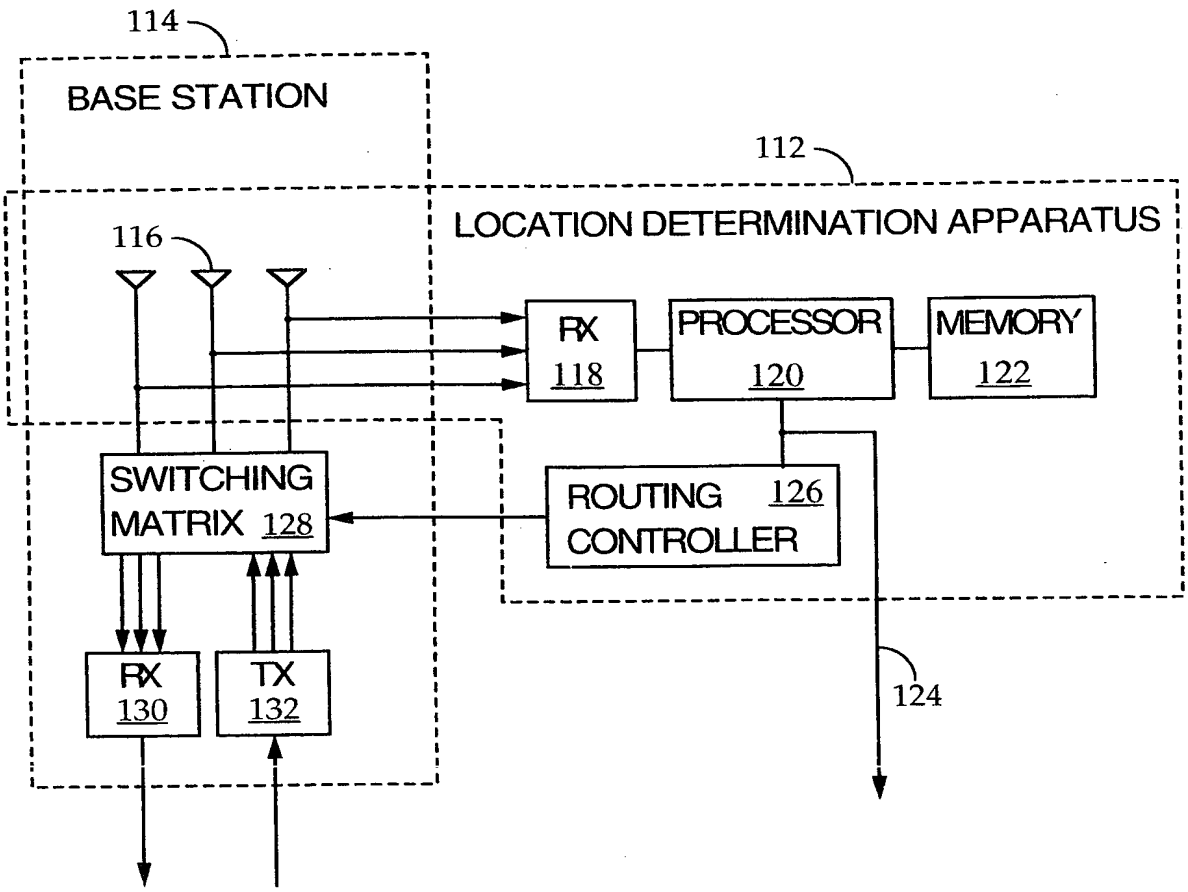

114
BASE STATION
112
LOCATION DETERMINATION APPARATUS
116
RX
118
PROCESSOR
120
MEMORY
122
SWITCHING
MATRIX 128
ROUTING 126
CONTROLLER
RX
130
TX
132
124

FIG. 2

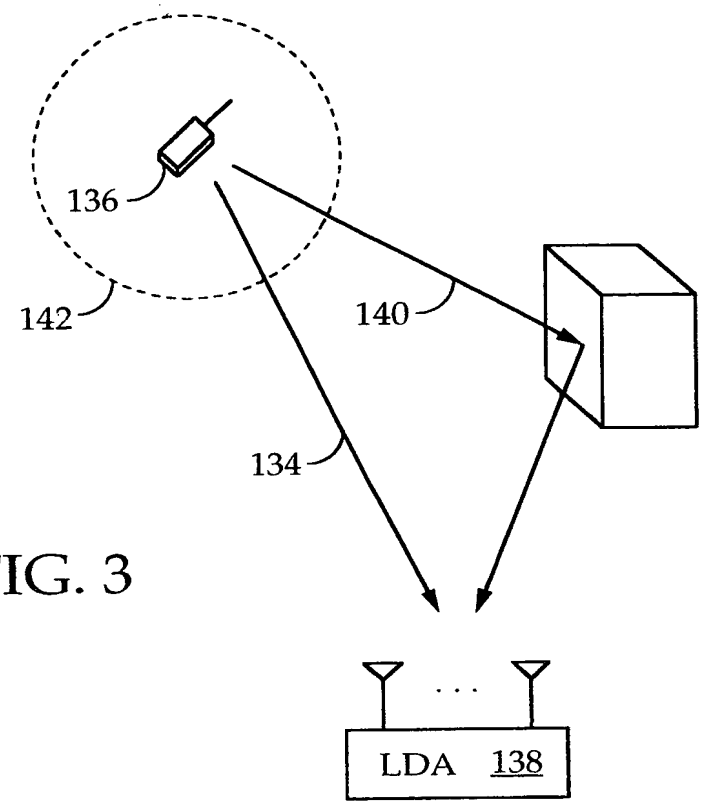


FIG. 3

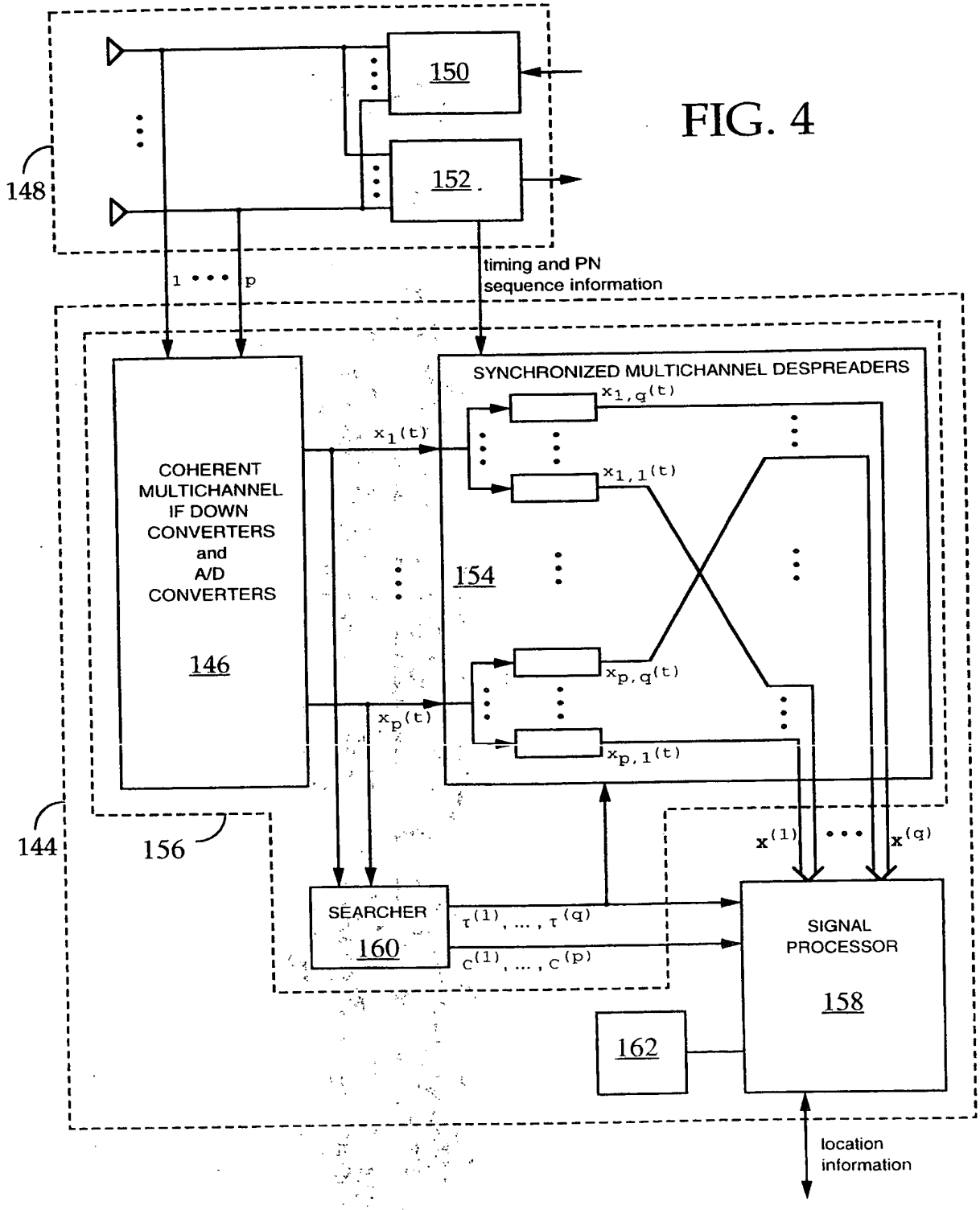
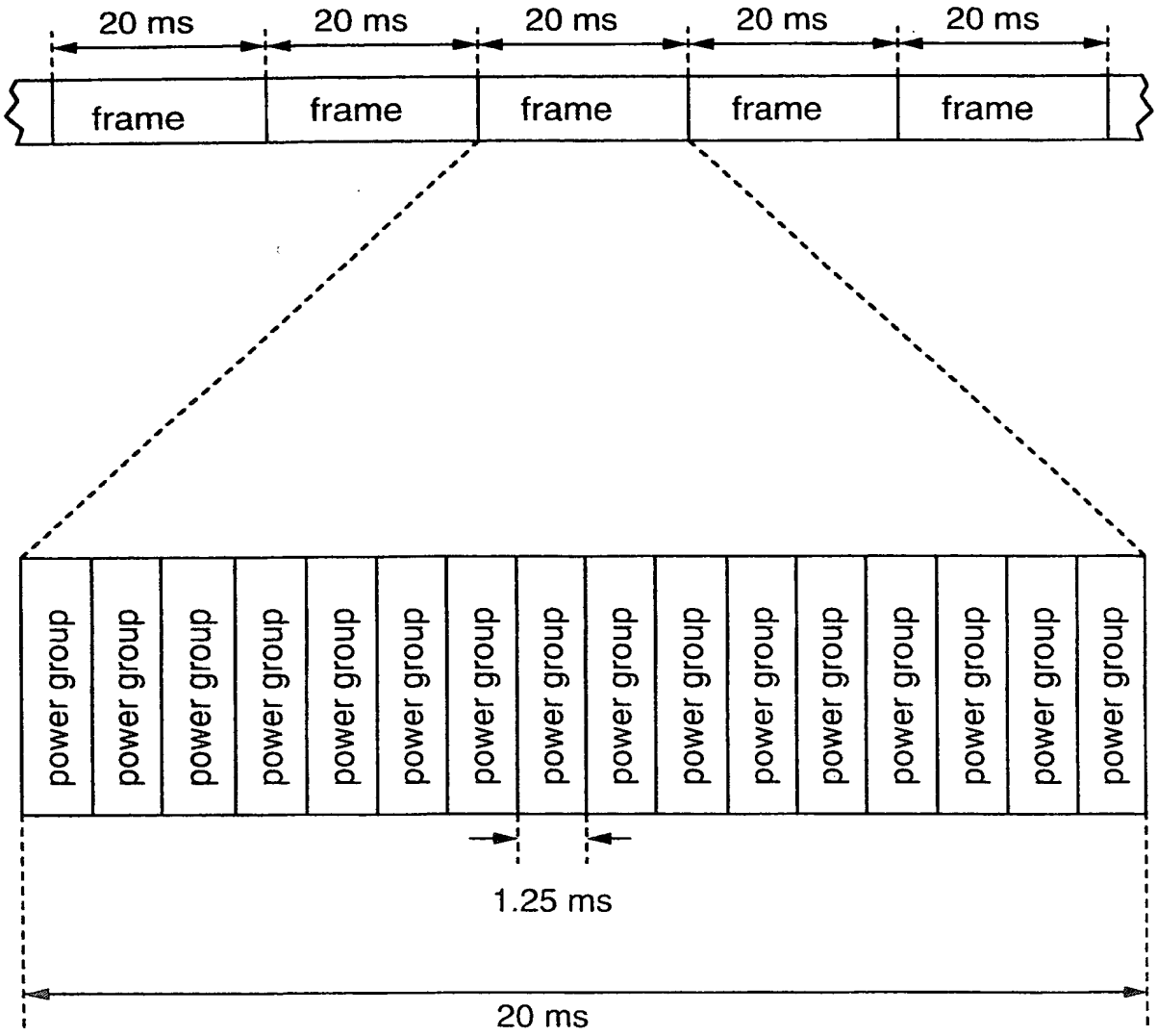

150
152
148
1 • • • p
timing and PN sequence information
SYNCHRONIZED MULTICHANNEL DESPREADERS
COHERENT MULTICHANNEL IF DOWN CONVERTERS and A/D CONVERTERS
146
154
144
156
SEARCHER
160
SIGNAL PROCESSOR
158
162
location information

FIG. 4

FIG. 5


20 ms
20 ms
20 ms
20 ms
20 ms
frame
frame
frame
frame
frame
power group
1.25 ms
20 ms

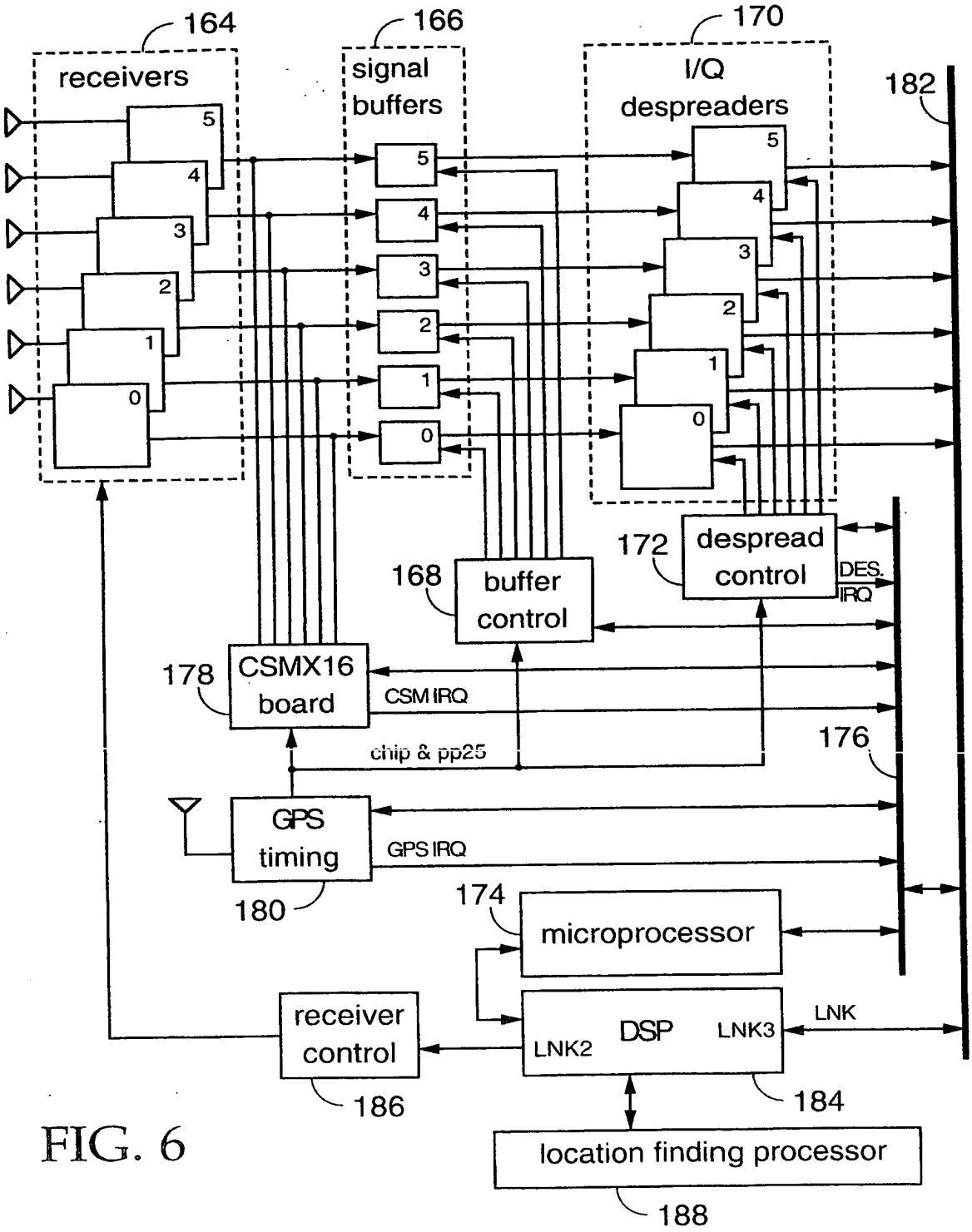

164
receivers
166
signal buffers
170
I/Q despreaders
182
5
4
3
2
1
0
172
despread control
DES. IRQ
168
buffer control
178
CSMX16 board
CSM IRQ
chip & pp25
176
GPS timing
GPS IRQ
180
174
microprocessor
receiver control
DSP
LNK2
LNK3
LNK
186
184
location finding processor
188

FIG. 6

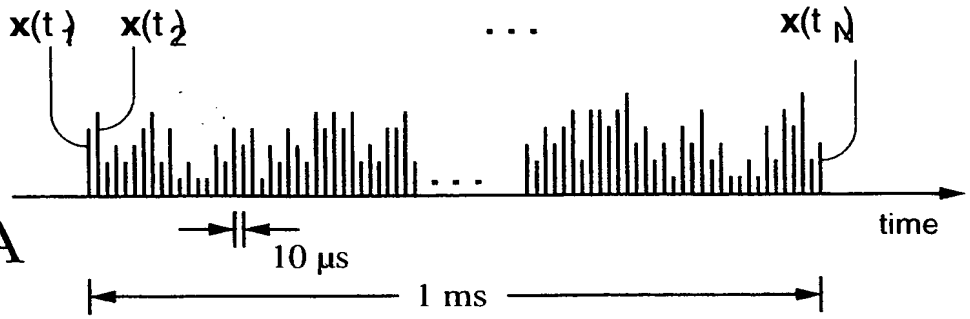


FIG. 7A

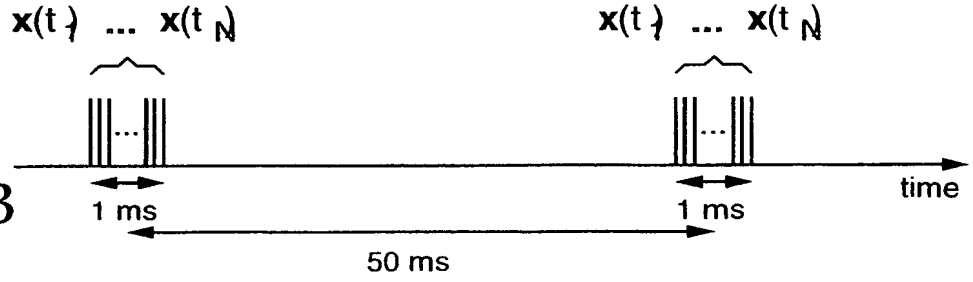


FIG. 7B

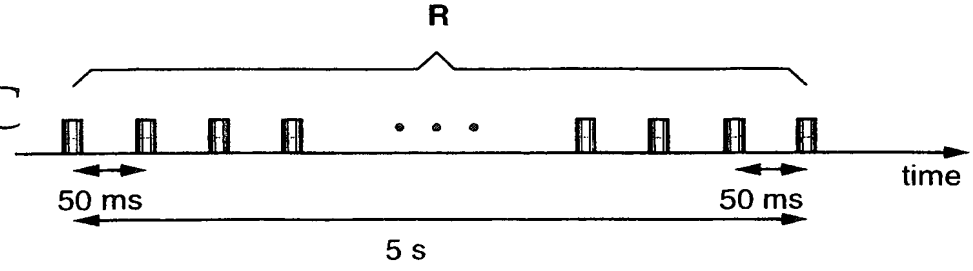


FIG. 7C

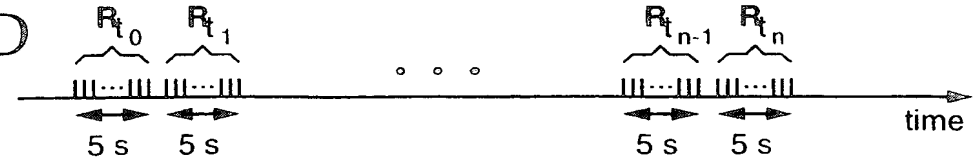


FIG. 7D

FIG. 8A

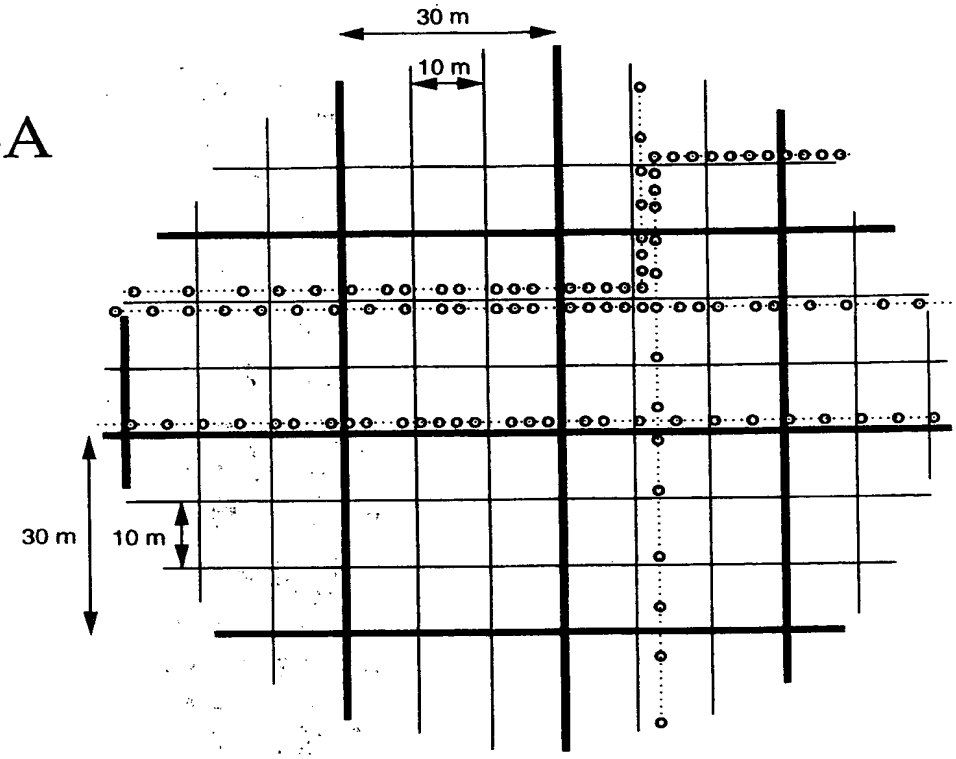
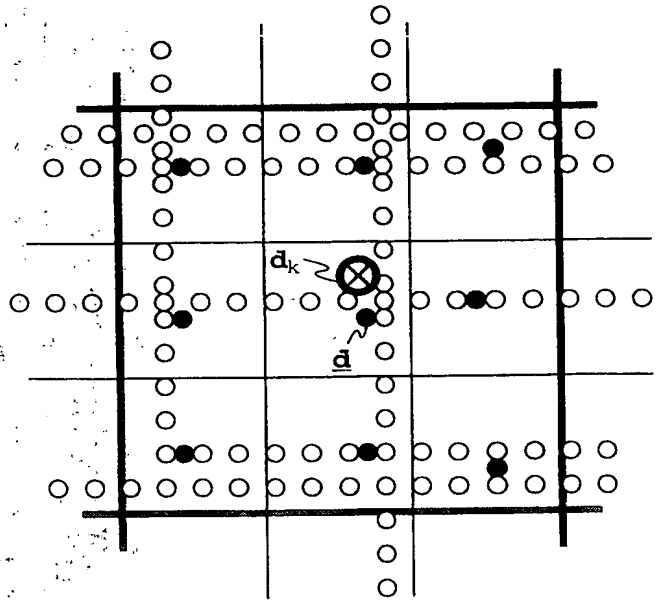


FIG. 8B



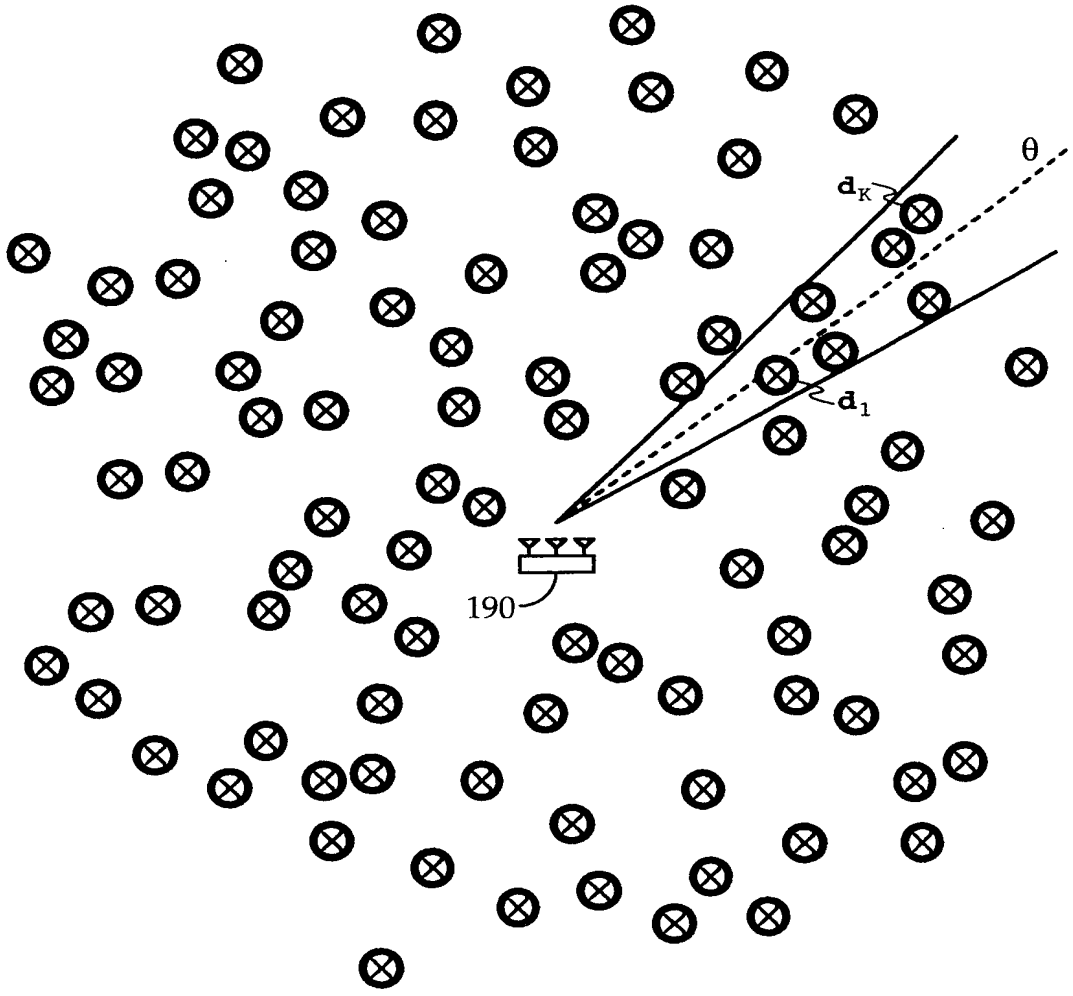


FIG. 9

FIG. 10

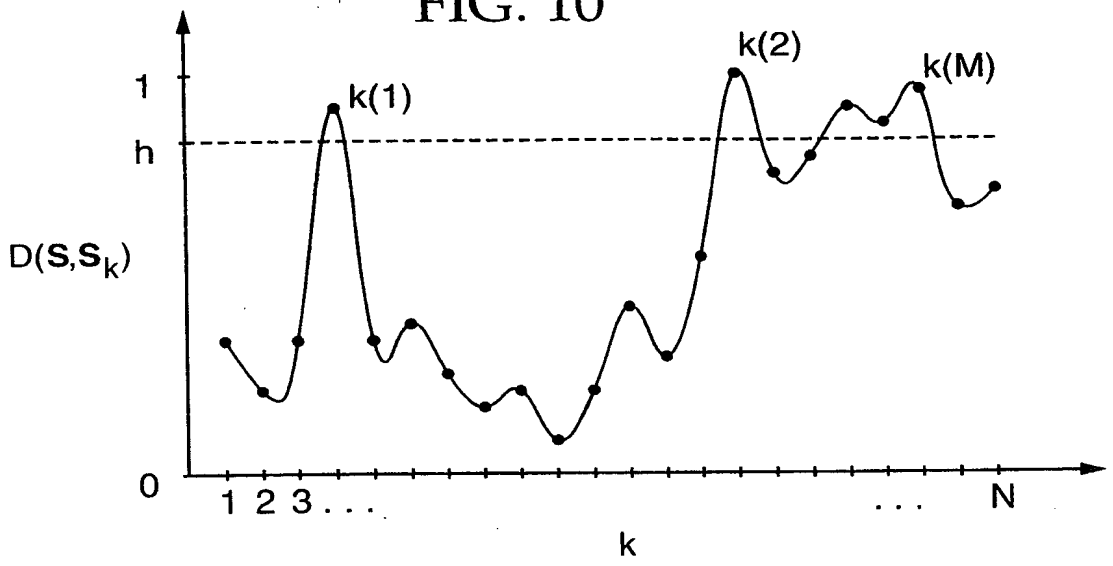
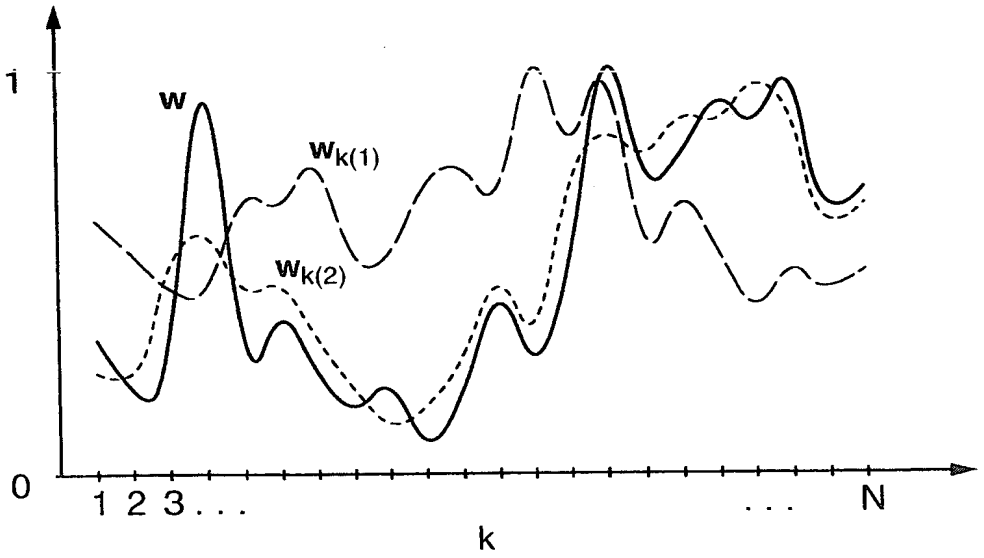


FIG. 11



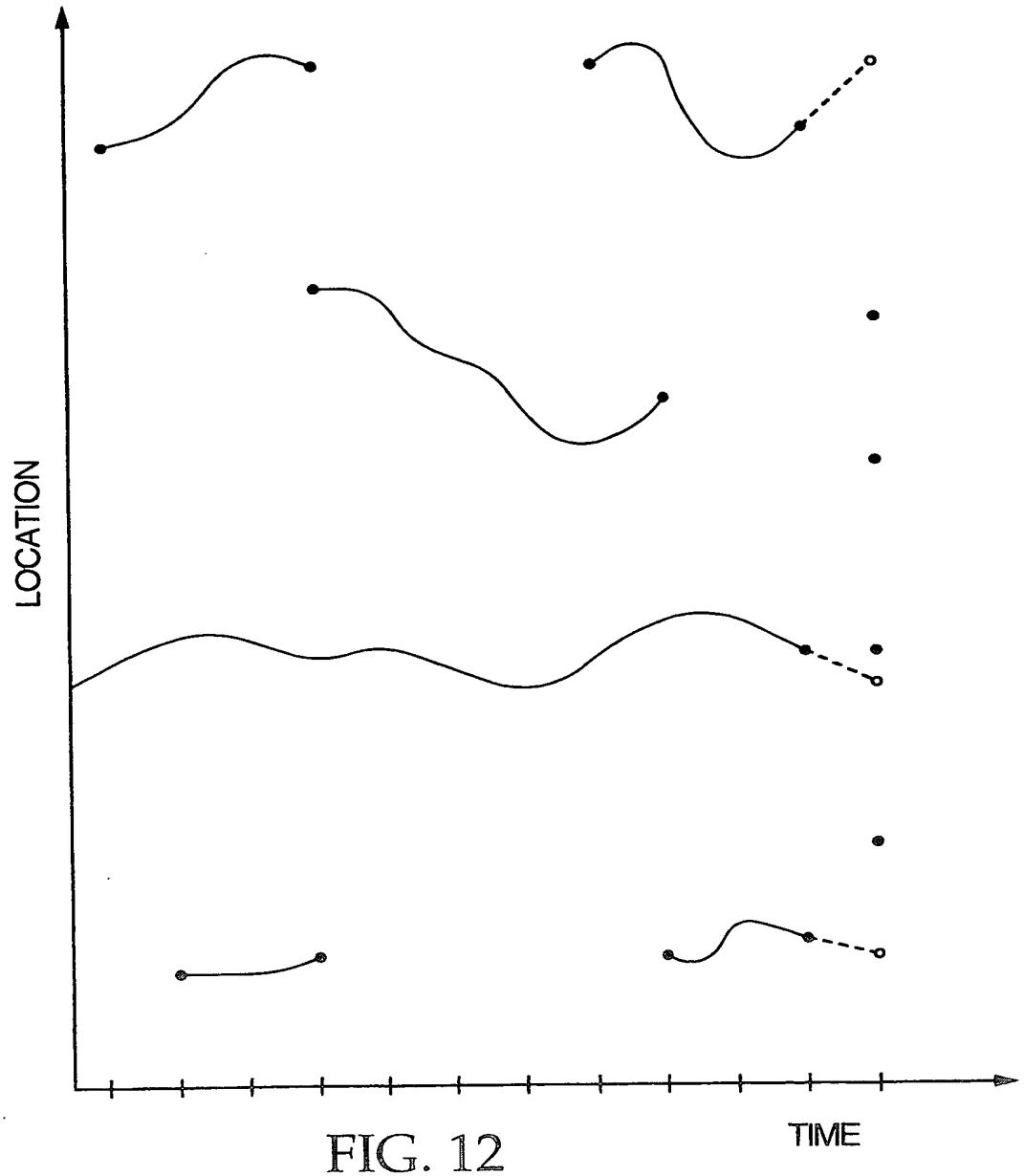


FIG. 12

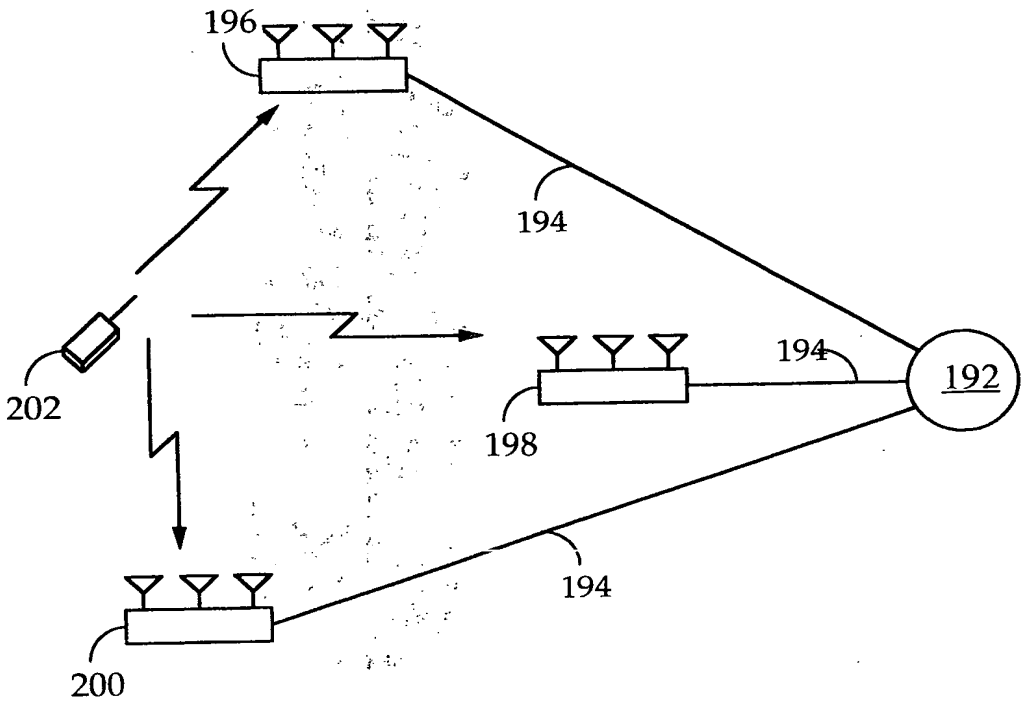


FIG. 13

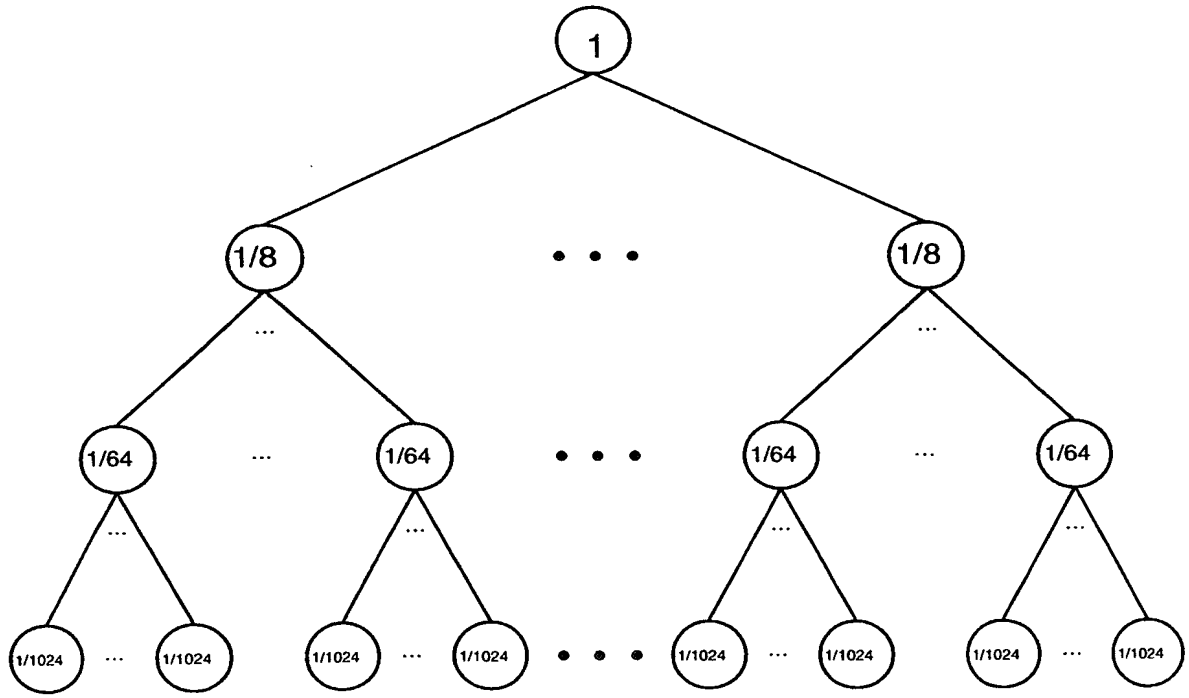

1
1/8
1/8
1/64
1/64
1/64
1/64
1/1024
1/1024
1/1024
1/1024
1/1024
1/1024
1/1024
1/1024

FIG. 14

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/24493

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) : G01S 3/02
US CL : 342/457
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 342/457, 450, 451, 453; 455/456

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,717,406 A (SANDERFORD ET AL) 10 FEBRUARY 1998 (10/02/98) IN ITS ENTIRETY	1-24
X	US 5,394,158 A (CHIA) 28 FEBRUARY 1995 (28/02/95) IN ITS ENTIRETY	1-24
X	HASHEMI, Homayoun. "Pulse Ranging Radiolocation Technique and its Application to Channel Assignment in Digital Cellular Radio", 41st IEEE Vehicular Technology Conference, 1991, 675-680.	1-24
A	JENG, S.S. et al. "Measurements of Spatial Signature of an Antenna Array", 6th Int'l Symposium Personal Indoor and Mobile Radio communications. 1995. pp. 669-672	1-24

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 08 MARCH 2000	Date of mailing of the international search report 10 MAY 2000
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer GREGORY C. ISSING Telephone No. (703) 306-4156

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/24493

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

IEEE DATABASE
SEARCH TERMS:(MULTIPATH, SIGNAL SIGNATURE, FINGERPRINT) AND (GEOLOCATION, LOCATION, POSITION) AND (CELLULAR,MOBILE)
BRS
SEARCH TERMS:SIMILAR TO ABOVE

BNSDOCID. <WO__ _ __ 0034799A1_1 >